

TAIL PROPERTIES OF MULTIVARIATE ARCHIMEDEAN COPULAS

Martin Larsson

MASTER OF ADVANCED STUDIES IN FINANCE

ETH ZURICH AND UNIVERSITY OF ZURICH

Supervised by Dr. Johanna Nešlehová

August 2008

Acknowledgment

I would like to thank Dr. Johanna Nešlehová for introducing me to the theory of Archimedean copulas, and for indicating the direction of investigation taken in this thesis. I am very grateful that she agreed to be my supervisor. Her comments, remarks and ideas have been crucial for the final result. Thanks are also due to Prof. Paul Embrechts for his constant support and encouragement.

Contents

1	Introduction and preliminaries	3
2	Archimedean copulas	8
2.1	Definition and examples	8
2.2	d -monotonicity and the Williamson d -transform	9
2.3	ℓ_1 -norm symmetric distributions	11
2.4	Archimedean copulas: characterization and stochastic representation . . .	13
2.5	A supplementary result	16
3	Regular variation and extreme value theory	19
3.1	Regular variation	19
3.2	Extreme value theory	20
4	Tail properties of the Williamson d-transform	23
4.1	Regular variation to the right, strict case	24
4.2	The Gumbel domain of attraction	26
4.3	Regular variation to the right, non-strict case	29
4.4	Behavior at the origin	32
4.5	Conditions on the inverse generator	35
5	Coefficients of tail dependence	37
5.1	Coefficient of the first kind	37
5.2	Coefficient of the second kind	40
6	Extreme value limits of Archimedean copulas	42
6.1	Componentwise minima of ℓ_1 -norm symmetric distributions	42
6.2	Componentwise maxima of ℓ_1 -norm symmetric distributions	45
7	Archimedean threshold copulas	49
7.1	Lower threshold copulas	49
7.2	Upper threshold copulas	53

Chapter 1

Introduction and preliminaries

Most problems in engineering and applied mathematics are motivated by some real world situation where one wants to make quantitative statements about the behavior and evolution of some system. Examples are mechanical engines, chemical processes, weather systems, or financial markets. These systems are rarely completely deterministic (in the sense that they can be described exactly by a deterministic model), but exhibit some randomness in their behavior. Sometimes these effects are negligible, but often it is necessary to take them into account. In some applications the influence of the random effects dominate over the deterministic structure, and a model that accurately captures the stochastic properties of the system is indispensable. If in such a situation the randomness is induced by different sources, or if several quantities are influenced differently by one source, a central issue is how to describe the dependence between these quantities; that the system behaves randomly does not mean that all components are independent.

Dependence between random variables can be very complicated, and not only in high dimensions. For example, an important but difficult issue is to find a relevant measure of the *strength* of dependence. It is easy to find examples where a very common such measure, the correlation coefficient, does not give any information at all¹.

Different approaches have been taken to study the concept of dependence, and a variety of results exist in probability theory. In this thesis we shall focus on so-called *copulas*, in particular a certain class of copulas called *Archimedean copulas*. A copula is an object that allows to isolate the dependence structure from the marginal behavior of a random vector, which probably is the main reason why the concept is theoretically useful. It also turns out that arbitrary marginal distributions can be combined to a multivariate joint distribution using any specified copula. This makes copulas a useful

¹Let X and Y be independent random variables, suppose $\mathbb{E}[Y] = 0$, and set $Z := XY$. Then $\text{Cov}[X, Z] = \mathbb{E}[XZ] - \mathbb{E}[X]\mathbb{E}[Z] = \mathbb{E}[X^2]\mathbb{E}[Y] - \mathbb{E}^2[X]\mathbb{E}[Y]$ by independence. The covariance is thus zero, but X and Z are clearly far from independent. See Embrechts et al. [9] for more on the intricacies of the correlation coefficient.

tool in practical applications, where it can be relatively easy to model the marginal distributions, but much harder to get an idea of the dependence. Using copulas, these two problems can be separated. Before continuing, we take the opportunity to mention that the copula approach to multivariate modeling is not uncontroversial. See Mikosch [21] for a discussion of the drawbacks.

What is then a copula? Formally it is defined as follows:

Definition 1.1 *A d -dimensional copula, generically denoted C , is a d -dimensional distribution function with standard uniform marginal distributions².*

It is not clear from the definition why copulas are useful for studying dependence. The following basic result due to Sklar [26] links the definition of copulas to their role in dependence modeling. We omit the proof, which is fairly simple in the case when the marginal distribution functions are continuous, and more involved when they are not.

Theorem 1.2 *Let H be a d -dimensional distribution function and F_1, \dots, F_d its one-dimensional marginals. Then there exists a copula C such that*

$$H(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)).$$

C is uniquely determined on $\text{Ran } F_1 \times \dots \times \text{Ran } F_d$, where $\text{Ran } F_i$ is the range of F_i . Conversely, given a copula C and univariate distribution functions F_1, \dots, F_d , the above relationship always defines a d -dimensional distribution function H .

Sklar's theorem clarifies the role of the copula C as carrier of the dependence structure, fully separated from the marginals. A result similar to Theorem 1.2 holds for survival functions. Although completely analogous we state it explicitly, since in our context survival functions will occur frequently.

Theorem 1.3 *Let \bar{H} be a d -dimensional survival function and $\bar{F}_1, \dots, \bar{F}_d$ the corresponding one-dimensional marginal survival functions. Then there exists a copula C such that*

$$\bar{H}(x_1, \dots, x_d) = C(\bar{F}_1(x_1), \dots, \bar{F}_d(x_d)).$$

The copula is uniquely determined on $\text{Ran } \bar{F}_1 \times \dots \times \text{Ran } \bar{F}_d$. Conversely, given a copula C and univariate survival functions $\bar{F}_1, \dots, \bar{F}_d$, the above relationship always defines a d -dimensional survival function \bar{H} .

The object C in Theorem 1.3 is called the *survival copula* of H (or \bar{H}). It is easy to show that if C_1 is the (ordinary) copula of H , and C_2 is its survival copula, then

$$(U_1, \dots, U_d) \sim C_1 \quad \text{if and only if} \quad (1 - U_1, \dots, 1 - U_d) \sim C_2,$$

where as usual the symbol \sim stands for “is distributed according to”. It is important to be clear about the distinction between a copula and the corresponding survival copula.

²Note that the condition on the marginals implies that the support of any d -dimensional copula is contained in the unit cube $[0, 1]^d$.

This sometimes becomes confusing, especially when the tails of a distribution are related to the tails of the copula (or survival copula). For example, the dependence in the upper tail of a distribution H is captured by the upper tail of its (ordinary) copula, but by the lower tail of its survival copula.

We note that Sklar's Theorem implies that given a joint distribution function H and its marginals F_1, \dots, F_d , the copula can be expressed as

$$C(u_1, \dots, u_d) = H(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d)). \quad (1.1)$$

Here F_i^{\leftarrow} is the *generalized inverse* of F_i , defined as

$$F_i^{\leftarrow}(u) := \inf\{x \in \mathbb{R} \mid F_i(x) \geq u\}, \quad u \in [0, 1].$$

Not all distribution functions F are invertible, but since they are non-decreasing the generalized inverse is always well-defined. The generalized inverse of the survival function \bar{F}_i is defined as $\bar{F}_i^{\leftarrow}(u) := F_i^{\leftarrow}(1 - u)$ for $u \in [0, 1]$. Similarly as in (1.1), the survival copula can be extracted from \bar{H} and $\bar{F}_1, \dots, \bar{F}_d$.

As already mentioned, this thesis is devoted to the class of Archimedean copulas. A lot of research has been done in this field, both concerning the general theory of copulas and Archimedean copulas in particular. Nelsen [22] contains a rigorous mathematical introduction, and McNeil, Frey and Embrechts [20] gives a good account of how to apply copulas in (financial) risk management. In this thesis we will exploit some recent advances in the understanding of the structure of such copulas. We give an introduction to Archimedean copulas and derive the necessary results and tools. This is carried out in Chapter 2. In particular our aim is to emphasize how these copulas arise as the dependence structures of certain mixture models.

The subsequent analysis will focus on the extremal behavior of Archimedean copulas. More precisely, we are interested in questions like how to describe the dependence between the largest observations (in a suitable sense) of a multivariate sample, or the dependence between observations exceeding some high threshold. Another issue is to what extent large outcomes occur simultaneously. All these concepts will be made precise as they are investigated. Some standard tools for dealing with such issues are introduced in Chapter 3, where we give a brief account of extreme value theory (EVT) and the theory of regularly varying functions.

Chapter 4, which contains the main theoretical contribution of this thesis, concerns the asymptotic behavior of a certain integral transform that proves to be of central importance in the theory of Archimedean copulas. The results in this section can, and should, also be interpreted as properties of the tails of certain Beta mixtures.

The last part of the thesis, Chapters 5–7, have the character of exercises in extreme value theory. It is shown how the results and concepts developed in Chapters 2–4 make it straightforward to determine extremal properties of Archimedean copulas, such

as coefficients of tail dependence, extreme value limits and limiting threshold copulas (notions that will be given precise definitions later on).

It is worth noting that we only are concerned with distributional properties of the random elements under considerations. We therefore only remark in passing that one always can find a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where all random elements live. This is a consequence of Skorokhod's Representation Theorem. That said, there will be no further mention of probability spaces.

Now some remarks on notation and terminology. Capital letters (X, Y, U, \dots) denote random variables. Vectors are written in boldface—so, for instance, $\mathbf{X} = (X_1, \dots, X_d)$ is a random vector taking values in \mathbb{R}^d , and $\mathbf{x} = (x_1, \dots, x_d)$ is a d -dimensional deterministic vector. The letter d will denote dimension, and is always an integer greater than or equal to two. Equalities, inequalities and arithmetic operations are to be interpreted componentwise. For example, $\mathbf{x} > \mathbf{y}$ means $x_i > y_i$ for all i .

With a slight abuse of language we will say that \bar{H} is a *survival function on the set* $A \subseteq \mathbb{R}^d$, when we mean that the corresponding random vector takes its values in A with probability one. In particular, saying that \bar{H} is a survival function on $\mathbb{R}_+^d := \{\mathbf{x} \in \mathbb{R}^d \mid \mathbf{x} \geq \mathbf{0}\}$ means that a random vector $\mathbf{X} \sim \bar{H}$ satisfies $\mathbf{X} \in \mathbb{R}_+^d$ a.s. Saying that \bar{H} is a survival function on the *interior* of \mathbb{R}_+^d means the same thing, except that we now exclude probability mass on the boundary of \mathbb{R}_+^d . Note that such statements about \bar{H} do *not* mean that the function $\mathbf{x} \mapsto \bar{H}(\mathbf{x})$ has its support confined to \mathbb{R}_+^d or its interior. They do, however, imply that \bar{H} is completely determined by its values in this set.

Several results, particularly in the last chapters, will concern *weak convergence* of random variables, denoted $X_n \Rightarrow X$ as $n \rightarrow \infty$. This means that the corresponding distribution functions F_n of X_n converge weakly to the distribution function F of X , i.e. that $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ for all x where the limit F is continuous. Weak convergence of random vectors, $\mathbf{X}_n \Rightarrow \mathbf{X}$, is defined analogously.

Finally, since distributions of mixtures, i.e. products of two random variables, will be a central theme, conditioning arguments are crucial. This can be quite a delicate issue (see e.g. Proschan and Presnel [24] for good illustrations of some of the pitfalls), which we nevertheless can circumvent using the following result:

Lemma 1.4 *Suppose that the \mathbb{R}^n -valued random variable \mathbf{X} and the \mathbb{R}^m -valued random variable \mathbf{Y} are independent, and let $h : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ be a measurable function such that $f(\mathbf{x}) := \mathbb{E}[h(\mathbf{x}, \mathbf{Y})]$ is well-defined for all $\mathbf{x} \in \mathbb{R}^n$. Then*

$$\mathbb{E}[h(\mathbf{X}, \mathbf{Y}) | \mathbf{X}] = f(\mathbf{X}).$$

In particular, if X is scalar with distribution function F , then

$$\mathbb{P}[XY > z] = \int_{\mathbb{R}} \mathbb{P}[Y > t^{-1}z] dF(t).$$

Proof. The first part is Corollary 4.38 in Breiman [4]. For the last part, let $h(x, \mathbf{y}) := \mathbb{1}_{\{\mathbf{y} > x^{-1}\mathbf{z}\}}$ and note that $\mathbb{P}[X\mathbf{Y} > \mathbf{z}] = \mathbb{E}[\mathbb{P}[\mathbf{Y} > X^{-1}\mathbf{z}|X]] = \mathbb{E}[\mathbb{E}[h(X, \mathbf{Y})|X]]$. Since $f(x) = \mathbb{P}[\mathbf{Y} > x^{-1}\mathbf{z}]$, the result follows from the first part of the lemma. \square

Chapter 2

Archimedean copulas

Archimedean copulas originally emerged as a family of copulas characterized by certain algebraic properties. This chapter is devoted to investigating their structure and characteristics. Section 2.1 defines Archimedean copulas and the so-called *Archimedean generators*. Some important examples are given. We then continue with some theoretical background on *d-monotone functions* (Section 2.2) and *ℓ_1 -norm symmetric distributions* (Section 2.3). This theory is necessary to prove (and understand) the main result, Theorem 2.12, which is given in Section 2.4. This result shows that Archimedean copulas can be *characterized* as the dependence structure of the very well-behaved family of ℓ_1 -norm symmetric distributions. It also characterizes the relevant class of Archimedean generators. Further we discuss interpretations of the Archimedean dependence structure. Section 2.5 concludes with a result that clarifies the correspondence between Archimedean copulas, their generators and ℓ_1 -norm symmetric distributions. The theory presented in this chapter (except Section 2.5) is based on McNeil and Nešlehová [19].

2.1 Definition and examples

Definition 2.1 (*Archimedean copula and Archimedean generator*) A d -dimensional copula C is called Archimedean if it has the representation

$$C(\mathbf{u}) = \psi(\psi^{-1}(u_1) + \dots + \psi^{-1}(u_d)), \quad \mathbf{u} \in \mathbb{R}_+^d,$$

where $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is an Archimedean generator, i.e. continuous, strictly decreasing on $\{\psi > 0\}$ and satisfying $\psi(0) = 1$ and $\lim_{x \rightarrow \infty} \psi(x) = 0$. Here ψ^{-1} is defined as $\psi^{-1}(u) := \inf\{x > 0 | \psi(x) = u\}$.

If $\psi^{-1}(0) = \infty$ the generator is said to be strict, otherwise non-strict.

The function ψ^{-1} is sometimes called the *pseudo-inverse* of ψ . Due to monotonicity and continuity of ψ , it coincides with the usual inverse of the restriction of ψ to $\{\psi > 0\}$. If

ψ is strict, this set is the whole of \mathbb{R}_+ . Non-strict generators are not invertible, but ψ^{-1} always is a right inverse of ψ , i.e. $\psi \circ \psi^{-1}(u) = u$ for all $u \in [0, 1]$.

Many authors define Archimedean copulas in terms of $\phi = \psi^{-1}$, and refer to this function as the generator. Although at first sight it is just a matter of convention, our definition turns out to be much more natural for studying the structure of these copulas. In this thesis, ψ is thus the generator, and $\phi = \psi^{-1}$ is referred to as the *inverse generator*.

An important observation is that ψ defines a survival function on \mathbb{R}_+ . This will be used later on.

Example 2.2 *The copulas $\Pi(\mathbf{u}) := u_1 u_2 \cdots u_d$ (the independence copula) and $W(u_1, u_2) := (u_1 + u_2 - 1)_+$ (the countermonotonicity copula) are both Archimedean, with generators $\psi_\Pi(x) = e^{-x}$ and $\psi_W(x) = (1 - x)_+$, respectively.*

The copulas $C_\theta^{\text{Cl}}(\mathbf{u})$ generated by $\psi_\theta^{\text{Cl}}(x) = (1 + \theta x)^{-1/\theta}$, $\theta \geq -1$ (the Clayton copula), and $C_\theta^{\text{Gu}}(\mathbf{u})$ generated by $\psi_\theta^{\text{Gu}}(x) = \exp\{-x^{1/\theta}\}$, $\theta \geq 1$ (the Gumbel copula), are two examples of parametric families of Archimedean copulas. They are among the most commonly encountered, and turn out to be particularly important in our context.

It is easy to check that the functions ψ_Π , ψ_W , ψ_θ^{Cl} and ψ_θ^{Gu} are all Archimedean generators. As for W (in the bivariate case) and Π (in any dimension), we even know that these functions are copulas¹. But we have not checked that the parametric families of claimed copulas are indeed copulas in the sense of Definition 1.1. To do this we need conditions for when an Archimedean generator ψ actually generates an Archimedean copula in a given dimension. In Section 2.4 we will characterize these generators and derive a very useful stochastic representation for Archimedean copulas. First, however, we need some background.

2.2 d -monotonicity and the Williamson d -transform

In the paper [27], R. E. Williamson studies the properties of real-valued functions having the following regularity property:

Definition 2.3 (*d -monotonicity*) *A function $f : [0, \infty) \rightarrow \mathbb{R}$ is called d -monotone for $d \geq 2$ if it is $d - 2$ times differentiable in $(0, \infty)$ and the derivatives satisfy*

$$(-1)^k f^{(k)}(x) \geq 0, \quad k = 0, 1, \dots, d - 2,$$

and $(-1)^{d-2} f^{(d-2)}$ is non-increasing and convex in $(0, \infty)$. It is said to be completely monotone if the above holds for all $d \geq 2$.

¹The marginals are clearly standard uniform. Π is a product of univariate distribution functions and hence itself a distribution function (with independent marginals). For W , observe that $W(u_1, u_2) = \mathbb{P}[U \leq u_1, 1 - U \leq u_2]$ for U standard uniform.

Note that although $f^{(d-2)}$ not necessarily has a derivative, it is convex and hence absolutely continuous on $(0, \infty)$. For ease of notation, we will write $f^{(d-1)}$ for the *right-continuous version* of the density of $f^{(d-2)}$.

Williamson [27] gives a representation theorem for d -monotone functions in terms of a certain integral transform, whose basic properties he also derives. These results were cast into the form presented here by McNeil and Nešlehová [19].

Definition 2.4 (*Williamson d -transform*) Let X be a non-negative random variable with distribution function F and $d \geq 2$ an integer. The Williamson d -transform of X (or F) is a real function on $[0, \infty)$ given by

$$\mathfrak{W}_d F(x) = \int_x^\infty \left(1 - \frac{x}{t}\right)^{d-1} dF(t) = \begin{cases} \mathbb{E} \left[\left(1 - \frac{x}{X}\right)_+^{d-1} \right] & \text{if } x > 0, \\ 1 - F(0) & \text{if } x = 0. \end{cases}.$$

Proposition 2.5 (*Properties of the Williamson d -transform*) Let F be a distribution function on \mathbb{R}_+ . The Williamson d -transform $f = \mathfrak{W}_d F$ satisfies the following properties:

- (i) (*Right continuity*) $f(x) \rightarrow f(0) = 1 - F(0)$ as $x \downarrow 0$.
- (ii) (*Order of decay*) $x^k f^{(k)}(x) \rightarrow 0$ as $x \rightarrow \infty$ for $k = 0, 1, \dots, d-1$.
- (iii) (*Invertibility*) The Williamson d -transform is invertible, and

$$F(x) = \mathfrak{W}_d^{-1} f(x) = 1 - \sum_{k=0}^{d-1} \frac{(-1)^k}{k!} x^k f^{(k)}(x), \quad x \geq 0.$$

Proof. Part (i) follows from the Dominated Convergence Theorem upon noting that

$$\left(1 - \frac{x}{X}\right)_+^{d-1} = \left(1 - \frac{x}{X}\right)^{d-1} \mathbb{1}_{\{X > x\}} \rightarrow \mathbb{1}_{\{X > 0\}} \text{ a.s.,} \quad x \downarrow 0.$$

Part (ii) is shown by Williamson [27], Lemma 1 (ii), and part (iii) is Proposition 5 (ii) in McNeil and Nešlehová [19]. \square

The following result is the main reason for studying the Williamson d -transform in connection with Archimedean copulas. Its relevance will be apparent later on.

Theorem 2.6 (*Representation theorem*) Let f be a real-valued function on \mathbb{R}_+ . Then f is d -monotone with $f(0) = p \in [0, 1]$ and $\lim_{x \rightarrow \infty} f(x) = 0$ if and only if $f = \mathfrak{W}_d F$ for some distribution function F on \mathbb{R}_+ .

Proof. The proof can be found in McNeil and Nešlehová [19], Proposition 5 (i). It is essentially a reformulation of Williamson's original proof in terms of distribution functions. \square

In Williamson's paper, the d -transform is viewed in the context of real analysis as a mapping on a set of non-decreasing functions on the positive half-line. With the formulation

here we view it in a probabilistic sense as a mapping on the set of distributions supported on the positive half-line. Using Lemma 1.4 (for the second equality below), this probabilistic interpretation can be taken further by noting that, for $G(x) = (1 - x)_+^{d-1}$,

$$\mathfrak{W}_d F(x) = \int_0^\infty G(x/t) dF(t) = \mathbb{E}[\mathbb{P}[X > x/Y | Y]] = \mathbb{P}[XY > x],$$

where X and Y are independent, $X \sim F$, and $Y \sim G$ and hence follows a Beta(1, $d - 1$) distribution. The Williamson d -transform thus describes the relationship between non-negative random variables X and certain Beta mixtures XY .

We conclude this section with a result that is *the* key to analyzing the structure of Archimedean copulas.

Proposition 2.7 *Let f be a real-valued function on $[0, \infty)$ and \bar{H} given by*

$$\bar{H}(\mathbf{x}) = f(x_1 + \dots + x_d), \quad \mathbf{x} \in \mathbb{R}_+^d.$$

Then \bar{H} is a survival function on the interior of \mathbb{R}_+^d if and only if f is d -monotone and satisfies $f(0) = 1$ and $\lim_{x \rightarrow \infty} f(x) = 0$.

Proof. See Proposition 2 in McNeil and Nešlehová [19]. The proof consists in showing that d -monotonicity is a necessary and sufficient condition for the (possibly signed) measure induced by \bar{H} to put non-negative mass on so-called *rectangle sets*, i.e. sets of the form $(a_1, b_1] \times \dots \times (a_d, b_d] \subset \mathbb{R}_+^d$. Together with conditions on \bar{H} and f at the origin and infinity, this is equivalent to \bar{H} inducing a probability measure. Note also that in our terminology, \bar{H} being a survival function on the interior of \mathbb{R}_+^d excludes it from assigning mass to the boundary of this set. \square

2.3 ℓ_1 -norm symmetric distributions

Although the relatively simple algebraic form (see Definition 2.1) is one of the reasons why Archimedean copulas are popular in some situations, this form does not help to interpret the corresponding dependence structure. Moreover, in many cases computations still become involved and seldom lead to any deeper understanding of *why* results hold true. This is in particular true in high dimensions, where calculations become cumbersome. Instead of working directly with the copulas, we would therefore prefer, if possible, to work with some suitable class of distributions having a simple structure. The family of ℓ_1 -norm symmetric distributions introduced in this section turns out to be an appropriate class. In what follows, \mathcal{S}^d will denote the d -dimensional unit simplex:

$$\mathcal{S}^d := \{\mathbf{x} \in \mathbb{R}_+^d \mid \|\mathbf{x}\|_1 = 1\}, \quad \text{where} \quad \|\mathbf{x}\|_1 = \sum_{i=1}^d |x_i|.$$

Definition 2.8 A random vector \mathbf{X} on \mathbb{R}_+^d is called ℓ_1 -norm symmetric if there exists a non-negative random variable R and a random vector \mathbf{S} , independent of R and uniformly distributed on the d -dimensional simplex \mathcal{S}^d , such that

$$\mathbf{X} \stackrel{d}{=} R\mathbf{S}.$$

The random variable R is referred to as the radial part of \mathbf{X} and its distribution as the radial distribution.

The stochastic representation in terms of a certain simplex mixture makes ℓ_1 -norm symmetric distributions very tractable objects. In the following we restrict attention to the case where R places no mass at the origin, i.e. $R > 0$ a.s.

Proposition 2.9 Suppose $\mathbf{X} \stackrel{d}{=} R\mathbf{S}$ is ℓ_1 -norm symmetric with $R > 0$ a.s. Then

- (i) $R \stackrel{d}{=} \|\mathbf{X}\|_1$ and $\mathbf{S} \stackrel{d}{=} \mathbf{X}/\|\mathbf{X}\|_1$
- (ii) $\|\mathbf{X}\|_1$ and $\mathbf{X}/\|\mathbf{X}\|_1$ are independent.

Proof. The proof relies on the following fact: if $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a measurable map and $\mathbf{Y} \stackrel{d}{=} \mathbf{Z}$ are identically distributed \mathbb{R}^n -valued random variables, then $f(\mathbf{Y}) \stackrel{d}{=} f(\mathbf{Z})$, see [11], p. 13. Letting $f : \mathbf{y} \mapsto (\|\mathbf{y}\|_1, \mathbf{y}/\|\mathbf{y}\|_1)$ for $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ we get

$$\left(\|\mathbf{X}\|_1, \frac{\mathbf{X}}{\|\mathbf{X}\|_1} \right) = f(\mathbf{X}) \stackrel{d}{=} f(R\mathbf{S}) = (R, \mathbf{S}),$$

using that $\|\mathbf{X}\|_1 > 0$ a.s. Now (i) clearly follows, and since we applied the result to the joint distribution of (R, \mathbf{S}) , we also get (ii). \square

The following characterization of ℓ_1 -norm symmetric distributions without point mass at the origin will be crucial in the sequel. Note that the proof of the direct half, (i) implies (ii), uses results on d -monotone functions. This fundamental connection between d -monotonicity and ℓ_1 -norm symmetric distributions is also evident when comparing the next result with Proposition 2.7.

Proposition 2.10 The following are equivalent:

- (i) $\bar{H}(\mathbf{x}) = f(x_1 + \dots + x_d)$, $\mathbf{x} \in \mathbb{R}_+^d$, is a survival function on the interior of \mathbb{R}_+^d .
- (ii) \bar{H} is the survival function of the ℓ_1 -norm symmetric random vector $\mathbf{X} \stackrel{d}{=} R\mathbf{S}$, $R > 0$ a.s., and $f(x) = \mathbb{P}[RS_1 > x] = \mathbb{P}[X_1 > x]$.

Proof. (ii) \implies (i): Let F be the distribution function of R . By Lemma 1.4 we have

$$\mathbb{P}[\mathbf{X} > \mathbf{x}] = \mathbb{E} \left[\mathbb{P} \left[\mathbf{S} > \frac{\mathbf{x}}{R} \middle| R \right] \right] = \int_{\mathbb{R}_+} \mathbb{P} \left[\mathbf{S} > \frac{\mathbf{x}}{t} \right] dF(t), \quad \mathbf{x} \in \mathbb{R}_+^d,$$

and similarly

$$f(x) = \mathbb{P}[RS_1 > x] = \int_{\mathbb{R}_+^d} \mathbb{P} \left[S_1 > \frac{x}{t} \right] dF(t), \quad x \geq 0.$$

It is well-known (see e.g. Fang et al. [11], Theorem 5.2 (i)) that the joint survival function of $\mathbf{S} \sim \text{unif}(\mathcal{S}^d)$ is given by $\mathbb{P}[\mathbf{S} > \mathbf{x}] = (1 - \|\mathbf{x}\|_1)_+^{d-1}$, $\mathbf{x} \in \mathbb{R}_+^d$. Hence $\mathbb{P}[S_1 > x] = (1 - x)_+^{d-1}$, $x \geq 0$, and we get

$$\mathbb{P}[\mathbf{S} > \mathbf{x}] = \mathbb{P}[S_1 > \|\mathbf{x}\|_1], \quad \mathbf{x} \in \mathbb{R}_+^d.$$

This finally implies that

$$\mathbb{P}[\mathbf{X} > \mathbf{x}] = \int_{\mathbb{R}_+} \mathbb{P}\left[S_1 > \frac{\|\mathbf{x}\|_1}{t}\right] dF(t) = f(\|\mathbf{x}\|_1), \quad \mathbf{x} \in \mathbb{R}_+^d.$$

(i) \implies (ii): Since (i) holds by assumption, f has to be d -monotone by Proposition 2.7. The representation theorem for d -monotone functions, Theorem 2.6, then gives a random variable R with distribution function F such that $f = \mathfrak{W}_d F$. Due to the invertibility of the Williamson d -transform (Proposition 2.5), F is uniquely determined by f , and by the discussion after the proof of Theorem 2.6, $\mathfrak{W}_d F(x) = \mathbb{P}[RS_1 > x]$ since we know that $S_1 \sim \text{Beta}(1, d-1)$. Finally, take \mathbf{S} uniform on \mathcal{S}^d and independent of R , and define $\mathbf{X} := R\mathbf{S}$. As in the proof of the first part of the theorem it follows that \bar{H} is the survival function of \mathbf{X} . \square

2.4 Archimedean copulas: characterization and stochastic representation

After the excursions in the two previous sections we now return to the Archimedean copulas and the problem of their characterization and representation. We will see how the theory of d -monotone functions and ℓ_1 -norm symmetric distributions enters the picture. The connection is actually little more than a simple observation; we use Sklar's theorem together with the fact that an Archimedean generator ψ is a survival function on \mathbb{R}_+ .

Suppose C is an Archimedean copula generated by ψ . If ψ is strict, the equality

$$\bar{H}(\mathbf{x}) := C(\psi(x_1), \dots, \psi(x_d)) = \psi(x_1 + \dots + x_d), \quad \mathbf{x} \in \mathbb{R}_+^d$$

certainly holds. If ψ is non-strict, the equality holds if $x_i \leq \psi^{-1}(0)$ for all i , since in this case $\psi^{-1} \circ \psi(x_i) = x_i$. If on the other hand $x_i > \psi^{-1}(0)$ for some i , the right side is clearly zero, as is the left side since $\psi(x_i) = 0$ and C is a distribution function with uniform marginals (hence no mass on the boundary of $[0, 1]^d$). So the equality holds also in this case. Sklar's theorem now implies that \bar{H} is a joint survival function on \mathbb{R}_+^d , and even on the interior of \mathbb{R}_+^d since $\psi(0) = 1$ by the definition of Archimedean generators.

Conversely, if a d -dimensional survival function \bar{H} on the interior of \mathbb{R}_+^d has the form

$$\bar{H}(\mathbf{x}) = \psi(x_1 + \dots + x_d), \quad \mathbf{x} \in \mathbb{R}_+^d$$

for some function ψ , the marginal survival functions are clearly $\psi(x)$, and Sklar's theorem gives the form of its survival copula C as

$$C(\mathbf{u}) = \psi(\psi^{\leftarrow}(u_1) + \dots + \psi^{\leftarrow}(u_d)).$$

If in addition ψ is an Archimedean generator we have $\psi^{\leftarrow} = \psi^{-1}$, and C is Archimedean by definition. We formulate these conclusions as

Lemma 2.11 *The function $C : [0, 1]^d \rightarrow [0, 1]$ given by*

$$C(\mathbf{u}) = \psi(\psi^{-1}(u_1) + \dots + \psi^{-1}(u_d)), \quad \mathbf{u} \in [0, 1]^d \quad (2.1)$$

is a d -dimensional Archimedean copula if and only if

$$\bar{H}(\mathbf{x}) = \psi(x_1 + \dots + x_d) = \psi(\|\mathbf{x}\|_1), \quad \mathbf{x} \in \mathbb{R}_+^d \quad (2.2)$$

is a joint survival function on the interior of \mathbb{R}_+^d . In this case, C is the survival copula of \bar{H} .

Proof. That C is the survival copula of \bar{H} is immediate from the discussion above. The equivalence assertion also follows as soon as we prove that the assumption (2.2) on \bar{H} implies that ψ is an Archimedean generator. Clearly $\psi(x) \rightarrow 0$ as $x \rightarrow \infty$ since \bar{H} is a survival function. Moreover, $\psi(0) = 1$ since \bar{H} is a survival function on the interior of \mathbb{R}_+^d . To deduce continuity and strict monotonicity we must invoke Proposition 2.7, which says that ψ is d -monotone. In particular, ψ is convex and hence continuous on $(0, \infty)$, and since (2.2) implies that ψ is a marginal survival function, it must be right continuous everywhere which excludes a jump at 0. Finally, since $\psi(x) \rightarrow 0$ as $x \rightarrow \infty$, convexity implies that ψ is strictly decreasing on the set where it is positive. \square

It is clear from Lemma 2.11 that if ψ generates a d -dimensional copula, it also generates k -dimensional copulas for all $k \leq d$. Indeed, $\psi(x_1 + \dots + x_k)$ is the k -dimensional marginal of $\bar{H}(\mathbf{x}) = \psi(x_1 + \dots + x_d)$, and we may apply Lemma 2.11 to this function.

We are now in a position to put the pieces together and give the main result that describes the structure of the Archimedean copulas. It is clear that the common denominator is the class of distributions whose survival functions can be written as a function of the ℓ_1 -norm of the argument. Due to Proposition 2.10, these are precisely the ℓ_1 -norm symmetric distributions without a point mass at the origin.

Theorem 2.12 *(Main result) The following are equivalent:*

- (i) $C(\mathbf{u}) = \psi(\psi^{-1}(u_1) + \dots + \psi^{-1}(u_d))$ is a d -dimensional Archimedean copula.
- (ii) ψ is a d -monotone Archimedean generator.
- (iii) $\bar{H}(\mathbf{x}) = \psi(x_1 + \dots + x_d)$, $\mathbf{x} \in \mathbb{R}_+^d$, is a survival function on the interior of \mathbb{R}_+^d .
- (iv) \bar{H} is the survival function of the ℓ_1 -norm symmetric random vector $\mathbf{X} \stackrel{d}{=} R\mathbf{S}$, $R > 0$ a.s., and $\psi(x) = \mathbb{P}[RS_1 > x] = \mathbb{P}[X_1 > x]$.

If (i)–(iv) hold, C is the survival copula corresponding to the survival function \bar{H} , and $\psi = \mathfrak{W}_d F$, where F is the distribution function of R .

Proof. The equivalence (i) \iff (iii) as well as the assertion that C is the survival copula of \bar{H} follow from Lemma 2.11. Moreover, (ii) \iff (iii) and (iii) \iff (iv) follow from Proposition 2.7 and Proposition 2.10, respectively. Finally, since S_1 is Beta(1, $d-1$) distributed, $\psi = \mathbb{P}[RS_1 > x]$ implies that $\psi = \mathfrak{W}_d F$ as in the discussion after the proof of Theorem 2.6. \square

Corollary 2.13 *Let \mathbf{U} be distributed according to the d -dimensional Archimedean copula C with generator ψ . Then $\mathbf{X} := (\psi^{-1}(U_1), \dots, \psi^{-1}(U_d))$ is ℓ_1 -norm symmetric with radial part $R \sim F$ satisfying $R > 0$ a.s. and $\psi = \mathfrak{W}_d F$.*

Proof. The survival function of \bar{H} of $\mathbf{X} := (\psi^{-1}(U_1), \dots, \psi^{-1}(U_d))$ is

$$\bar{H}(\mathbf{x}) = \mathbb{P}[\psi^{-1}(U_1) > x_1, \dots, \psi^{-1}(U_d) > x_d] = \mathbb{P}[U_1 \leq \psi(x_1), \dots, U_d \leq \psi(x_d)],$$

since ψ is decreasing. \mathbf{U} is distributed according to C , so

$$\bar{H}(\mathbf{x}) = \psi(\psi^{-1}(\psi(x_1)) + \dots + \psi^{-1}(\psi(x_1))) = \psi(\|\mathbf{x}\|_1), \quad \mathbf{x} \in \mathbb{R}_+^d,$$

where the last equality follows as in the proof of Lemma 2.11. The assertion now follows from Theorem 2.12. \square

Part (ii) of Theorem 2.12 gives a necessary and sufficient condition that can be used to check whether a given Archimedean generator indeed generates a d -dimensional copula. One can for instance show that this is the case for the Gumbel generator $\psi_\theta^{\text{Gu}}(x)$ in Example 2.2 if $\theta \geq 1$. The Clayton generator $\psi_\theta^{\text{Cl}}(x)$ turns out to be d -monotone for $\theta \geq -1/(d-1)$, see McNeil and Nešlehová [19], Example 4. Part (iv) gives a stochastic representation that greatly simplifies computations and can be used for simulation purposes. It is also the key to important interpretations of the Archimedean dependence structures, which is essential when justifying the use of these copulas in practice.

1. We know that an ℓ_1 -norm symmetric random variable $\mathbf{X} \stackrel{d}{=} R\mathbf{S}$ has the properties that $\|\mathbf{X}\|_1 \stackrel{d}{=} R$ and $\mathbf{X}/\|\mathbf{X}\|_1 \stackrel{d}{=} \mathbf{S}$. Hence it can be viewed as an allocation of a random total amount R of some resource among d units, such that S_i is the fraction allocated to unit i .
2. Moreover, $\|\mathbf{X}\|_1$ and $\mathbf{X}/\|\mathbf{X}\|_1$ are independent, meaning that the individual fractions do not depend on the total allocated amount.
3. Finally, requiring uniformity of \mathbf{S} on the unit simplex \mathcal{S}^d amounts to saying that all possible combinations of allocated fractions are equally likely to occur. This clearly imposes a strong symmetry on the distribution.

Archimedean copulas are the natural dependence structures of such “fair-share” distributions. One application from finance is in credit risk, where large portfolios of credit

deals (e.g. loans) are commonplace. In many cases such portfolios, or sub-portfolios, possess a large degree of homogeneity, which is consistent with point 3. Viewing the total loss as the “resource” (in credit risk it is common practice only to consider losses; R will thus be positive) and assuming that the size of the loss does not influence its distribution between clients, we have 1. and 2. The latter is reasonable only if the client exposures are of comparable orders of magnitude; knowing that the total loss was large, one could otherwise deduce that this large loss probably was caused by defaults among high-exposure clients rather than low-exposure clients. This would contradict independence between $\|\mathbf{X}\|_1$ and $\mathbf{X}/\|\mathbf{X}\|_1$.

2.5 A supplementary result

Due to the invertibility of the Williamson d -transform, Theorem 2.12 implies that there is a one-to-one correspondence between a.s. strictly positive random variables, or, equivalently, ℓ_1 -norm symmetric distributions without a point mass at the origin, and d -monotone Archimedean generators. The correspondence between the generators and the Archimedean copulas is, however, *not* one-to-one. Indeed, $\psi(x)$ and $\tilde{\psi}(x) := \psi(cx)$, $c > 0$, give rise to the same copula. The corresponding radial distributions are $R = \mathfrak{W}_d^{-1}\psi$ and $\tilde{R} := R/c$. The next result shows that for a fixed copula C , the possible generators (or radial parts) cannot be more different than that.

Proposition 2.14 *(i) Two Archimedean generators ψ and $\tilde{\psi}$ generate the same copula if and only if $\tilde{\psi}(cx) = \psi(x)$ for some $c > 0$.*

(ii) Let \mathbf{X} and $\tilde{\mathbf{X}}$ be ℓ_1 -norm symmetrically distributed. They have identical survival copulas if and only if their radial parts R and \tilde{R} satisfy $\tilde{R} \stackrel{d}{=} R/c$ for some $c > 0$.

Proof. Due to the one-to-one correspondence between the generators and radial parts as discussed above, it suffices to prove one of the two statements. We give a proof of (i). Sufficiency follows immediately. For necessity, suppose ψ and $\tilde{\psi}$ generate the same copula. Then

$$\tilde{\psi}(\tilde{\psi}^{-1}(u_1) + \cdots + \tilde{\psi}^{-1}(u_d)) = \psi(\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)), \quad u_i \in [0, 1].$$

Now define $T := \tilde{\psi}^{-1} \circ \psi$. Since the pseudo-inverse always is a right inverse, we have $T \circ \psi^{-1} = \tilde{\psi}^{-1}$ and $\tilde{\psi} \circ T = \psi$. The above equality is thus equivalent to

$$\tilde{\psi}(T \circ \psi^{-1}(u_1) + \cdots + T \circ \psi^{-1}(u_d)) = \tilde{\psi} \circ T(\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)),$$

or, with $x_i := \psi^{-1}(u_i) \in [0, x^*]$, where $x^* := \psi^{-1}(0)$,

$$\tilde{\psi}(T(x_1) + \cdots + T(x_d)) = \tilde{\psi} \circ T(x_1 + \cdots + x_d), \quad x_i \in [0, x^*]. \quad (2.3)$$

We would now like to apply $\tilde{\psi}^{-1}$ on both sides of (2.3) to get a functional equation for T , but $\tilde{\psi}$ is only invertible on $\{\tilde{\psi} > 0\}$. Now, if $x_1 + \cdots + x_d < x^*$ then $T(x_1 + \cdots + x_d) <$

$\tilde{\psi}^{-1}(0)$, implying that the right side of (2.3) is strictly positive. The left side is then also strictly positive, and since $x_1 + \cdots + x_d < x^*$ implies that $x_i < x^*$ for each i (all x_i are positive) we conclude

$$T(x_1) + \cdots + T(x_d) = T(x_1 + \cdots + x_d), \quad x_i > 0, \ x_1 + \cdots + x_d < x^*.$$

By continuity of T this holds also for $x_i \geq 0, x_1 + \cdots + x_d < x^*$. With $x_1 = \cdots = x_d = 0$ we get $dT(0) = T(0)$, so $T(0) = 0$. Moreover, setting $x_3 = \cdots = x_d = 0$ and using that $T(0) = 0$ yields

$$T(x_1) + T(x_2) = T(x_1 + x_2), \quad x_1, x_2 \geq 0, \ x_1 + x_2 < x^*.$$

By Lemma 2.15 below, the only solutions to this functional equation are $T(x) = cx$, $x \in [0, x^*]$, where $c \in \mathbb{R}$. Hence $\psi(x) = \tilde{\psi}(cx)$ for $x \in [0, x^*]$, and since both ψ and $\tilde{\psi}$ are Archimedean generators it is immediate that $c > 0$. If $x^* = \infty$ we are done. Otherwise we note that $\psi(x^*) = 0$, and so $\tilde{\psi}(x) = \psi(x/c) = 0$ for $x \geq cx^*$, since $\tilde{\psi}$ is non-increasing. \square

Lemma 2.15 *Let f be a real-valued, continuous function on \mathbb{R}_+ such that²*

$$f(x) + f(y) = f(x + y), \quad x, y \geq 0, \ x + y < a, \quad (2.4)$$

where $a \in (0, \infty]$. Then $f(x) = cx$ for all $x \in [0, a]$ and some $c \in \mathbb{R}$.

Proof. We prove the lemma in the case where $a < \infty$. If $a = \infty$ the proof is simpler in the sense that one does not have to be as careful to avoid going outside $[0, a]$, where the functional equation is defined. Otherwise it is very similar.

We may without loss of generality assume that $a = 1$. Once this case is established, the general case follows by considering $\tilde{f}(x) := f(ax)$. Moreover, by continuity of f we have (2.4) for $x, y \geq 0, x + y \leq a$. Now assume that

$$(i) \ f(nx) = nf(x) \text{ for all } x \text{ such that } 0 \leq nx \leq 1$$

holds for some $n \in \mathbb{N}$, and take x such that $0 \leq (n+1)x \leq 1$. Then

$$f(nx + x) = f(nx) + f(x) = nf(x) + f(x),$$

using (2.4) and (i). The latter is trivially satisfied for $n = 1$, so by induction it holds for all n . Take now $m \in \mathbb{N}$ and $x \in [0, 1]$. Then $m\frac{x}{m} \in [0, 1]$, so (i) yields $f(x) = f(m\frac{x}{m}) = mf(\frac{x}{m})$, or, equivalently,

$$(ii) \ f(\frac{1}{m}x) = \frac{1}{m}f(x) \text{ for all } x \in [0, 1].$$

²Note the similarity to the well-known Cauchy functional equation: $f(x) + f(y) = f(x + y)$ for all $x \in \mathbb{R}$. See Bingham et al. [2], Section 1.1, for a treatment. However, although similar, our case requires a separate proof.

Finally take $x \in [0, 1]$ rational. Then $x = \frac{n}{m}$ for some $n \in \mathbb{N} \cup \{0\}$ and $m \in \mathbb{N}$ with $m \geq n$. Since $0 \leq n \frac{1}{m} \leq 1$ and $\frac{1}{m} \leq 1$, we may apply (i) and (ii) to obtain

$$f(x) = f\left(\frac{n}{m}\right) = nf\left(\frac{1}{m}\right) = \frac{n}{m}f(1) = xf(1).$$

With $c := f(1)$ we thus have $f(x) = cx$ for all *rational* $x \in [0, 1]$, and by density and continuity it also holds for all $x \in [0, 1]$: take $x_n \rightarrow x \in [0, 1]$ with $x_n \in [0, 1]$ rational, and deduce that $f(x) = \lim_{n \rightarrow \infty} f(x_n) = cx$. \square

Chapter 3

Regular variation and extreme value theory

3.1 Regular variation

We summarize here some basic results from the theory of real-valued regularly varying functions. The main reference for this theory is Bingham et al. [2], and the results are stated (and occasionally proven) in many other places as well. Two good references are Embrechts et al. [10] and Resnick [25]. The latter reference also discusses so-called multivariate regular variation in the context of multivariate extreme value theory.

We use the following notation: for two positive functions f and g we write $f(x) \sim g(x)$, $x \rightarrow \infty$, if $\lim_{x \rightarrow \infty} f(x)/g(x) = 1$. Whenever $f(x) \sim 0$ appears, it is to be understood as $\lim_{x \rightarrow \infty} f(x) = 0$. There will be no danger of confusing the two meanings of the symbol “ \sim ”.

Definition 3.1 *A positive, measurable function f on $(0, \infty)$ is called regularly varying with index ρ , written $f \in \text{RV}_\rho$, if for any $c > 0$,*

$$\lim_{x \rightarrow \infty} \frac{f(cx)}{f(x)} = c^\rho.$$

It is called rapidly varying (with index $-\infty$), $f \in \text{RV}_{-\infty}$, if

$$\lim_{x \rightarrow \infty} \frac{f(cx)}{f(x)} = \begin{cases} 0 & \text{if } c > 1 \\ \infty & \text{if } 0 < c < 1. \end{cases}$$

Functions in RV_0 are called slowly varying.

Note that a regularly varying function $f \in \text{RV}_\rho$ always can be represented as $f(x) = x^\rho L(x)$, with $L \in \text{RV}_0$; just set $L(x) := f(x)x^{-\rho}$. In this thesis, the function f will

usually be the tail of some distribution function F , i.e. $f = 1 - F = \bar{F}$. We will then sometimes write $X \in \text{RV}_\rho$ for a random variable $X \sim F$ to denote that $\bar{F} \in \text{RV}_\rho$.

Theorem 3.2 (*Karamata's theorem*) *Let $L \in \text{RV}_0$ be locally bounded in $[0, \infty]$. Then*

- (i) *If $\sigma < 0$, then $\int_x^\infty t^{\sigma-1} L(t) dt \sim -x^\sigma L(x)/\sigma$, $x \rightarrow \infty$.*
- (ii) *If $\sigma > 0$, then $\int_0^x t^{\sigma-1} L(t) dt \sim x^\sigma f(x)/\sigma$, $x \rightarrow \infty$.*
- (iii) *The function $x \mapsto \int_0^x t^{-1} L(t) dt$ is slowly varying, if finite.*

Proof. These are Theorem 1.5.10, 1.5.8 and 1.5.9a in Bingham et al. [2]. \square

Theorem 3.3 (*Monotone Density Theorem*) *Suppose $U(x) \sim cx^\rho L(x)$, $x \rightarrow \infty$, where $c \in \mathbb{R}$, $\rho \in \mathbb{R}$, $L \in \text{RV}_0$ and suppose u is some eventually monotone function.*

- (i) *If $U(x) = \int_x^\infty u(y) dy$, then $u(x) \sim -c\rho x^{\rho-1} L(x)$ as $x \rightarrow \infty$.*
- (ii) *If $U(x) = \int_0^x u(y) dy$, then $u(x) \sim c\rho x^{\rho-1} L(x)$ as $x \rightarrow \infty$.*

Proof. (ii) is Theorem 1.7.2 in Bingham et al. [2], and the proof of (i) is virtually identical. \square

Theorem 3.4 *Suppose $f \in \text{RV}_{-\infty}$ is non-increasing and take $\sigma \in \mathbb{R}$. Then $\int_{x_0}^\infty t^\sigma f(t) dt < \infty$ for some $x_0 \geq 0$, and*

$$\lim_{x \rightarrow \infty} \frac{\int_x^\infty t^\sigma f(t) dt}{x^{\sigma+1} f(x)} = 0.$$

Proof. This is (part of) Theorem A3.12 in Embrechts et al. [10]. \square

3.2 Extreme value theory

Univariate extreme value theory in its most basic form is concerned with the study of limit relations of the form

$$F^n(c_n x + d_n) \rightarrow G(x) \text{ weakly,} \quad n \rightarrow \infty, \quad (3.1)$$

where F and G are distribution functions with G non-degenerate, $c_n > 0$, and $d_n \in \mathbb{R}$. The probabilistic interpretation is the following: Suppose X_1, X_2, \dots are i.i.d. random variables distributed according to F , and let $M_n := \max(X_1, \dots, X_n)$ denote the largest of the first n variables in the sequence. Then, by independence, $\mathbb{P}[M_n \leq x] = \mathbb{P}^n[X_1 \leq x] = F^n(x)$, and after translation and scaling by d_n and c_n , $\mathbb{P}[c_n^{-1}(M_n - d_n) \leq x] = F^n(c_n x + d_n)$. The relation (3.1) thus expresses *weak convergence of the normalized maxima* $c_n^{-1}(M_n - d_n)$.

If (3.1) holds for some specific F and non-degenerate G , we say that F lies in the *maximum domain of attraction* (or just *domain of attraction*) of G , denoted $F \in \text{MDA}(G)$. The first issue to resolve is whether the limit G is unique. More precisely, is it possible to find normalizing sequences in (3.1) so that for the same F we get convergence to two different limits G_1 and G_2 ? It turns out that this can happen, but in that case G_1 and G_2 are of the same *type*, meaning that $G_1(x) = G_2(Ax + B)$ for some constants $A > 0$ and $B \in \mathbb{R}$. This is part of the Convergence to Types Theorem (see Proposition 0.2 in Resnick [25]).

The next task is now to determine the possible limits G . This was done by Fisher and Tippett [13], and it turns out that the limit (up to type) belongs to one of three families:

- (i) The Fréchet family: $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}$, $x \geq 0$, for some $\alpha \in (0, \infty)$.
- (ii) The Gumbel distribution: $\Lambda(x) = \exp\{-e^{-x}\}$, $x \in \mathbb{R}$.
- (iii) The Weibull family: $\Psi_\alpha(x) = \exp\{-(-x)^\alpha\}$, $x < 0$, for some $\alpha \in (0, \infty)$.

For each of these three families it now remains to characterize the respective domains of attraction and the corresponding normalizing sequences (c_n) and (d_n) .

Theorem 3.5 *Let F be a distribution function.*

- (i) (Fréchet case) $F \in \text{MDA}(\Phi_\alpha)$ if and only if $\bar{F} \in \text{RV}_{-\alpha}$ for $\alpha \in (0, \infty)$.
- (ii) (Gumbel case) $F \in \text{MDA}(\Lambda)$, if and only if there exists a positive function a , called the auxiliary function of F , with $a(x)/x \rightarrow 0$ as $x \rightarrow x^*$, where x^* is the right endpoint of F , such that

$$\lim_{x \uparrow x^*} \frac{\bar{F}(x + a(x)t)}{\bar{F}(x)} = e^{-t}, \quad \forall t \in \mathbb{R}. \quad (3.2)$$

In this case one may choose $a(x) := \int_x^{x^} \bar{F}(t) dt / \bar{F}(x)$. If F has a density, we may alternatively choose $a(x) := -\bar{F}(x) / \bar{F}'(x)$. It is always possible to find an auxiliary function that is absolutely continuous with $\lim_{x \rightarrow x^*} a'(x) = 0$.*

- (iii) (Weibull case) $F \in \text{MDA}(\Psi_\alpha)$ if and only if F has finite right endpoint x^* and $\bar{F}(x^* - x^{-1}) \in \text{RV}_{-\alpha}$ for $\alpha \in (0, \infty)$.

Proof. For (i) and (iii), see Gnedenko [16], Theorem 4 and Theorem 5, respectively. For the first part of (ii), see Gnedenko [16], Theorem 7. For the two choices of auxiliary function, see de Haan [17], Lemma 5 and Resnick [25], Proposition 1.1 (b). That the auxiliary function can be taken absolutely continuous with the given property can be deduced from Resnick [25], Proposition 1.4 and the proof of Proposition 1.1(a) in the same reference. \square

We note that in the Gumbel case, all auxiliary functions of a given distribution F are

tail equivalent. This convenient fact is proved in [25], p. 26¹. Moreover, if $F \in \text{MDA}(\Lambda)$ has unbounded support, then \bar{F} is rapidly varying, $\bar{F} \in \text{RV}_{-\infty}$.

In this thesis we will not need the full answer to the question of how to find appropriate normalizing sequences. For this we refer to Resnick [25], Proposition 1.11 and 1.13 for the Fréchet and Weibull domains, respectively, and Proposition 1.1, 1.4 and 1.19 for the Gumbel domain. For our purposes the following is sufficient:

Proposition 3.6 *If $F \in \text{MDA}(\Psi_\alpha)$, the normalizations $c_n = x^* - (1/\bar{F})^\leftarrow(n)$ and $d_n = x^*$ yield $G = \Psi_\alpha$ in (3.1). Here x^* is the right endpoint of F .*

Proof. This follows from Proposition 1.13 in Resnick [25], or can be deduced from the proof of Theorem 5 in Gnedenko's paper [16]. \square

¹Alternatively one may invert both sides of (3.2) to obtain $\{\bar{F}^\leftarrow(\bar{F}(x)t) - x\}/a(x) \rightarrow -\log t$ as $x \uparrow x^*$. If this also holds with $a(x)$ replaced by some $\hat{a}(x)$, then clearly $a(x)/\hat{a}(x) \rightarrow 1$. That inverting (3.2) is allowed can be proven in the same way as Proposition 0.1 in Resnick [25].

Chapter 4

Tail properties of the Williamson d -transform

We are interested in the extremal behavior of Archimedean copulas, which depends on the generator near its right endpoint and near the origin. We prefer, however, to work with ℓ_1 -norm symmetric distributions instead of Archimedean generators, and it is thus of interest to connect the extremal behavior of the radial distributions to that of the corresponding generators. By Theorem 2.12 this amounts to investigating the tail properties of the Williamson d -transform. This program is carried out in the present chapter. We separate the analysis into four different cases: regular variation at infinity, the Gumbel domain of attraction, the behavior near the right endpoint in the case of bounded support, and the behavior at the origin.

Due to the interpretation of the Williamson d -transform as the survival function of certain Beta mixtures, the proofs of some of the results below can be significantly simplified using the following result by Breiman, slightly extended to fit our needs:

Lemma 4.1 *Let X and Y be two non-negative random variables such that $\bar{F} \in \text{RV}_{-\alpha}$, $\alpha \in (0, \infty)$, where F is the distribution function of X , and $\mathbb{E}[Y^{\alpha+\varepsilon}] < \infty$ for some $\varepsilon > 0$. Then*

$$\frac{\mathbb{P}[XY > x]}{\mathbb{P}[X > x]} \rightarrow \mathbb{E}[Y^\alpha], \quad x \rightarrow \infty.$$

If $Y \in (0, 1)$ a.s., the conclusion holds also with $\alpha = 0$ and $\alpha = \infty$, where in the latter case $\mathbb{E}[Y^\alpha] = 0$ by convention.

Proof. The case when $\alpha \in (0, \infty)$ is covered by Breiman's original result [3], so we suppose that $Y \in (0, 1)$ a.s. Denoting by G the distribution function of Y we get, using Lemma 1.4,

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}[XY > x]}{\mathbb{P}[X > x]} = \lim_{x \rightarrow \infty} \int_0^1 \frac{\bar{F}(x/t)}{\bar{F}(x)} dG(t).$$

Since \bar{F} is non-increasing and $t \in (0, 1)$, the integrand is bounded by one, so we may interchange limit and integration by the Dominated Convergence Theorem. If $\alpha = 0$, $\bar{F}(x/t)/\bar{F}(x) \rightarrow 1$ as $x \rightarrow \infty$. If $\alpha = \infty$, the limit is zero. In either case we clearly reach the desired conclusion. \square

As an aid for the intuition, the reader is encouraged to think of the results in this chapter as probabilistic statements about mixture variables, and not only as analytical properties of integral transforms.

4.1 Regular variation to the right, strict case

A strict Archimedean generator ψ , viewed as a survival function, can be expected to possess interesting extremal behavior only if $\psi \in \text{RV}_{-\alpha}$ for some $\alpha \in (0, \infty]$; this follows from Theorem 3.5, since any other behavior of ψ would exclude it from being in some domain of attraction. By definition, $\psi \in \text{RV}_{-\alpha}$ means that

$$\lim_{x \rightarrow \infty} \frac{\psi(cx)}{\psi(x)} = c^{-\alpha}, \quad \forall c > 0. \quad (4.1)$$

Now, by Theorem 2.12, ψ has to be d -monotone in order to generate a copula in dimension d . At first sight, this seems to be a fairly strong regularity condition on ψ , especially for large d , and one wonders if perhaps (4.1) always holds, at least if we also allow $\alpha = 0$. This question is particularly relevant in view of the fact that all commonly encountered families of Archimedean copulas have generators that do satisfy (4.1). We now give a counterexample to show that this is not always the case.

Example 4.2 *Define*

$$\psi(x) := \frac{1 + a \sin(\log(1 + x))}{1 + x}, \quad x \geq 0,$$

for some $a > 0$ to be determined. Then $\psi(0) = 1$ and $\lim_{x \rightarrow \infty} \psi(x) = 0$, and if $a < 1$, then $\psi(x) > 0$ for all $x \geq 0$. We show that for any d , one can choose $a > 0$ such that ψ is d -monotone. Suppose that for some $k \in \{0, 1, \dots\}$,

$$\psi^{(k)}(x) = (-1)^k \frac{k! + ac_k \sin(\varphi_k + \log(1 + x))}{(1 + x)^{k+1}}, \quad x \geq 0, \quad (4.2)$$

where c_k and φ_k are constants. Differentiating and using standard trigonometric identities one obtains

$$\begin{aligned} \psi^{(k+1)}(x) &= (-1)^{k+1} \frac{(k+1)! + a(k+1)c_k \sin(\varphi_k + \log(1 + x)) - ac_k \cos(\varphi_k + \log(1 + x))}{(1 + x)^{k+2}} \\ &= (-1)^{k+1} \frac{(k+1)! + ac_{k+1} \sin(\varphi_{k+1} + \log(1 + x))}{(1 + x)^{k+2}}, \end{aligned}$$

where $c_{k+1} = c_k \sqrt{k^2 + 2k + 2}$ and $\varphi_{k+1} = \varphi_k + \arctan(1/(k+1))$. Since (4.2) holds for $k = 0$ with $c_0 = 1$ and $\varphi_0 = 0$, it follows by induction that it holds for all $k \in \{0, 1, \dots\}$. Since $\sin(y) \geq -1$ for all $y \in \mathbb{R}$, it is clear that given $d \geq 2$, $(-1)^k \psi^{(k)}(x) \geq 0$ for all $k = 1, \dots, d-1$, provided $0 < a \leq k!/c_k$ for each k . Since for $k \geq 2$ it holds that $c_k > k!$ and $\max(c_0, \dots, c_k) = c_k$, such an a can indeed be found and satisfies $a < 1$ (take $a = (d-1)!/c_{d-1}$), implying that ψ is indeed d -monotone.

We now consider the behavior of $\psi(cx)/\psi(x)$ as $x \rightarrow \infty$ along the sequence $x_n = e^{n\pi}$. We write

$$\frac{\psi(cx_n)}{\psi(x_n)} = \frac{1 + a \sin(\log(1 + cx_n))}{1 + a \sin(\log(1 + x_n))} \frac{1 + x_n}{1 + cx_n} =: A(x_n) \frac{1 + x_n}{1 + cx_n}.$$

Since $(1 + x_n)/(1 + cx_n) \rightarrow 1/c$, we focus attention on $A(x_n)$. With x_n as above we get

$$A(x_n) = \frac{1 + a \sin(n\pi + \log(e^{-n\pi} + c))}{1 + a \sin(n\pi + \log(e^{-n\pi} + 1))} = \frac{1 + a(-1)^n \sin(\log(e^{-n\pi} + c))}{1 + a(-1)^n \sin(\log(e^{-n\pi} + 1))}.$$

The denominator converges to one as $n \rightarrow \infty$, but the numerator does not converge in general: along the subsequence where all n are even, it converges to $1 + a \sin(\log c)$, while if all n are odd, the limit is $1 - a \sin(\log c)$. So unless $\log c = m\pi$ for some integer m , $\psi(cx_n)/\psi(x_n)$ does not converge.

We remark that there are known families of Archimedean copulas where the generators satisfy (4.1) with $\alpha = 0$. One such example is $\psi(x) = \theta / \log(e^\theta + x)$ with $\theta \in (0, \infty)$, which generates bivariate Archimedean copulas (see Nelsen [22], item 19 of Table 4.1, where however the inverse generator $\phi(t) = e^{\theta/t} - e^\theta$ is given instead of ψ). In such cases it is still possible to draw conclusions about the tail behavior of the copula, see Charpentiers and Segers [6], although the corresponding ℓ_1 -norm symmetric distributions do not possess non-degenerate extremal behavior (under affine normalizations).

After this introductory discussion, we turn to sufficient and in most cases necessary conditions on the radial part of the corresponding ℓ_1 -norm symmetric distribution for (4.1) to hold.

Proposition 4.3 *Let F be a distribution function with $F(0) = 0$ and let $\psi = \mathfrak{W}_d F$, $d \geq 2$, be its Williamson d -transform. Then*

- (i) *If $\bar{F} \in RV_{-\alpha}$, $\alpha \in [0, \infty)$, then $\psi \in RV_{-\alpha}$.*
- (ii) *If $\psi \in RV_{-\alpha}$, $\alpha \in (0, \infty)$, then $\bar{F} \in RV_{-\alpha}$.*

Proof. Suppose $R \sim F$ and let S_1 be Beta(1, $d-1$) distributed and independent of R . Then, as we have seen (e.g. in Theorem 2.12), $\psi(x) = \mathbb{P}[RS_1 > x]$. Observing that $S_1 \in (0, 1)$ a.s., we obtain from Lemma 4.1 that $\bar{F}(x) \sim B_1 \psi(x)$ as $x \rightarrow \infty$ for some $B_1 > 0$, whence the statement (i).

For part (ii), we have by assumption that $\psi(x) = x^{-\alpha} L(x)$ for some $\alpha > 0$ and slowly varying L . We first show by induction that, for $k = 0, 1, \dots, d-1$,

$$(-1)^k x^k \psi^{(k)}(x) \sim c_k x^{-\alpha} L(x) = c_k \psi(x), \quad x \rightarrow \infty,$$

where $c_0 = 1$ and $c_k = \prod_{j=0}^{k-1} (j + \alpha)$ for $k = 1, \dots, d-1$. For $k = 0$ the statement is trivially true. For $k = 0, \dots, d-2$,

$$\psi^{(k)}(x) = - \int_x^\infty \psi^{(k+1)}(y) dy,$$

since $\psi^{(k)}(x) \rightarrow 0$ as $x \rightarrow \infty$. Moreover, $\psi^{(k+1)}$ is monotone since ψ is d -monotone, and $\psi^{(k)}(x) \sim (-1)^k c_k x^{-k-\alpha} L(x)$ by the induction assumption. The Monotone Density Theorem now implies that $\psi^{(k+1)}(x) \sim (-1)^{k+1} c_{k+1} (-k-\alpha) x^{-k+1-\alpha} L(x)$, which is precisely the statement we wanted to prove.

Inserting this into the inverse Williamson d -transform yields

$$\frac{\bar{F}(x)}{\psi(x)} = \frac{1 - \mathfrak{W}_d^{-1} \psi(x)}{\psi(x)} = \sum_{k=0}^{d-1} \frac{1}{k!} \frac{(-1)^k x^k \psi^{(k)}(x)}{\psi(x)} \rightarrow \sum_{k=0}^{d-1} \frac{1}{k!} c_k, \quad x \rightarrow \infty.$$

The limit is clearly positive since $\alpha > 0$. □

Remark. One can show that

$$\sum_{k=0}^{d-1} \frac{1}{k!} c_k = 1 + \sum_{k=1}^{d-1} \frac{1}{k!} \prod_{j=0}^{k-1} (j + \alpha) = \prod_{k=1}^{d-1} \left(1 + \frac{\alpha}{k}\right),$$

which by Breiman's result is equal to $1/\mathbb{E}[S_1^\alpha]$. As a by-product we thus get an expression for the raw moments of Beta($1, d-1$) distributed random variables.

4.2 The Gumbel domain of attraction

If $\bar{F} \in \text{RV}_{-\infty}$, which for instance is the case if $F \in \text{MDA}(\Lambda)$ and has unbounded support, we may use Lemma 4.1 similarly as in the proof of Proposition 4.3 (i) to conclude that $\psi(x)/\bar{F}(x) \rightarrow 0$ as $x \rightarrow \infty$. This gives no information about the precise behavior of the tail of ψ . In what follows we derive more precise results in the case where F and/or ψ is in the Gumbel domain of attraction. This is the relevant case in the context of extreme value theory. It turns out to be unnecessary to distinguish between the strict and the non-strict cases.

We proceed by proving a number of intermediate results. The condition (3.2) from Theorem 3.5 will appear several times, and we emphasize that the requirements on the auxiliary function a are part of this condition. Namely, $a(x) > 0$ and $a(x)/x \rightarrow 0$ as x tends to the right endpoint x^* (which may be infinity).

Lemma 4.4 *Suppose the non-negative, non-increasing function f satisfies (3.2), i.e.*

$$\lim_{x \uparrow x^*} \frac{f(x + a(x)t)}{f(x)} = e^{-t}, \quad \forall t \in \mathbb{R},$$

where $a(x) > 0$ and $a(x)/x \rightarrow 0$ as $x \uparrow x^*$. If f has a density f' that is monotone in a neighborhood of x^* , then

$$\lim_{x \uparrow x^*} \frac{f'(x + a(x)t)}{f'(x)} = e^{-t}, \quad \forall t \in \mathbb{R},$$

Proof. The core of the argument is similar to the proof of Theorem 3.10.11 in Bingham et al. [2]. For all x large enough, $x + a(x)t = x(1 + ta(x)/x)$ is close to x^* for any $t \in \mathbb{R}$, since $a(x)/x \rightarrow 0$ as $x \uparrow x^*$. Now choose any $u, v \in \mathbb{R}$, $u > v$. For all x sufficiently close to x^* , say $x \in (x_0, x^*)$, we thus have that $x + a(x)v$ and $x + a(x)u$ are both in the region where f' is monotone. Since f is non-increasing, $-f'$ is non-negative. We look at the case where $-f'$ is non-increasing near x^* , the non-decreasing case being completely analogous. We obtain

$$\begin{aligned} -\frac{f(x + a(x)u) - f(x + a(x)v)}{f(x)} &= \frac{a(x)}{f(x)} \int_v^u -f'(x + a(x)t) dt \\ &\leq -\frac{a(x)}{f(x)} f'(x + a(x)v)(u - v), \end{aligned}$$

where each expression is non-negative. The left side tends to $e^{-v} - e^{-u}$ as $x \rightarrow x^*$, and hence

$$\liminf_{x \rightarrow x^*} -\frac{a(x)}{f(x)} f'(x + a(x)v) \geq \lim_{u \downarrow v} \frac{e^{-v} - e^{-u}}{u - v} = e^{-v},$$

where the last equality follows by e.g. l'Hôpital's rule. Similarly one shows

$$\limsup_{x \rightarrow x^*} -\frac{a(x)}{f(x)} f'(x + a(x)u) \leq e^{-u},$$

and so

$$\lim_{x \rightarrow x^*} -\frac{a(x)}{f(x)} f'(x + a(x)t) = e^{-t}, \quad \forall t \in \mathbb{R}. \quad (4.3)$$

In particular, we may take $t = 0$ to obtain $-f(x)/a(x) \sim f'(x)$, $x \rightarrow x^*$, which implies that f' is non-zero near x^* . Substituting back into (4.3) gives the conclusion. \square

Lemma 4.5 *Suppose the function f satisfies (3.2). Then, for any $\rho \in \mathbb{R}$, $g(x) = x^\rho f(x)$ also satisfies (3.2) with the same auxiliary function.*

Proof. A simple computation yields

$$\frac{g(x + a(x)t)}{g(x)} = \left(1 + \frac{a(x)}{x}t\right)^\rho \frac{f(x + a(x)t)}{f(x)} \rightarrow e^{-t},$$

since f satisfies (3.2), which in particular means that $a(x)/x \rightarrow 0$ as $x \rightarrow x^*$. \square

Lemma 4.6 *Suppose the non-negative, non-increasing function f satisfies (3.2). Then $g(x) = \int_x^\infty f(t)dt$ also satisfies (3.2) with the same auxiliary function.*

Proof. Suppose \hat{a} is the auxiliary function of f . By Theorem 3.5 we can find another auxiliary function a that is absolutely continuous with $a'(x) \rightarrow 0$, $x \rightarrow x^*$. L'Hôpital's rule then gives

$$\lim_{x \uparrow x^*} \frac{g(x + a(x)t)}{g(x)} = \lim_{x \uparrow x^*} \frac{-f(x + a(x)t)}{-f(x)} (1 + a'(x)t).$$

The limit on the right is clearly e^{-t} by the properties of f and a . So g satisfies (3.2) with auxiliary function a , hence also with \hat{a} , since the two are tail equivalent, i.e. $\hat{a}(x)/a(x) \rightarrow 1$, $x \rightarrow x^*$. \square

Lemma 4.7 *Let F be a distribution function with $F(0) = 0$ and let $\psi = \mathfrak{W}_d F$. Then $\psi^{(d-1)}$ is differentiable if and only if F is differentiable, and in this case*

$$\bar{F}'(x) = \frac{(-1)^{d-1}}{(d-1)!} x^{d-1} \psi^{(d)}(x).$$

Proof. From the inversion formula of the Williamson d -transform in Proposition 2.5, it is clear that $\psi^{(d-1)}$ is differentiable if and only if F is, and that the derivative \bar{F}' in this case is given by $\bar{F}'(x) = f_d(x)$, where we define

$$f_n(x) := D \left[\sum_{k=0}^{n-1} \frac{(-1)^k}{k!} x^k \psi^{(k)}(x) \right]$$

for all n such that $\psi^{(n)}$ exists. Here $D[\cdot]$ denotes differentiation. We are done if we can show that for any such n ,

$$f_n(x) = \frac{(-1)^{n-1}}{(n-1)!} x^{n-1} \psi^{(n)}(x). \quad (4.4)$$

We argue by induction over n . For $n = 2$, $f_n(x) = D[\psi(x) - x\psi'(x)] = -x\psi''(x)$, so (4.4) holds. Suppose (4.4) holds for $n-1$ and $\psi^{(n)}$ exists. Then

$$\begin{aligned} f_n(x) &= f_{n-1}(x) + D \left[\frac{(-1)^{n-1}}{(n-1)!} x^{n-1} \psi^{(n-1)}(x) \right] \\ &= \frac{(-1)^{n-1}}{(n-1)!} x^{n-1} \psi^{(n)}(x). \end{aligned}$$

\square

We are now ready to prove the following proposition concerning the connection between \bar{F} and ψ when either is in the Gumbel domain of attraction. Note the assumption that \bar{F} or ψ be differentiable—we are not treating the most general case here.

Proposition 4.8 *Let F be a distribution function with $F(0) = 0$, let $\psi = \mathfrak{W}_d F$ for $d \geq 2$ be its Williamson d -transform, and suppose that either F or $\psi^{(d-1)}$ is differentiable. Both are then differentiable, and if we in addition assume that F' is monotone in a neighborhood of its right endpoint x^* , then $F \in \text{MDA}(\Lambda)$ if and only if $\psi \in \text{MDA}(\Lambda)$. In this case,*

$$a(x) = \frac{\int_x^{x^*} \bar{F}(t) dt}{\bar{F}(x)} \quad \text{and} \quad a(x) = \frac{\int_x^{x^*} \psi(t) dt}{\psi(x)}$$

are possible choices of auxiliary function for both \bar{F} and ψ .

Proof. Lemma 4.7 implies that both \bar{F} and ψ are differentiable and satisfy

$$\bar{F}'(x) = \frac{(-1)^{d-1}}{(d-1)!} x^{d-1} \psi^{(d)}(x). \quad (4.5)$$

We start with necessity, so suppose \bar{F} is in the Gumbel domain. Theorem 3.5 and the monotonicity assumption on $F' = -\bar{F}'$ then enable us to use Lemma 4.4 to conclude that $-\bar{F}'$, and hence $(-1)^d x^{d-1} \psi^{(d)}$, satisfies (3.2). The auxiliary function carries over, and Lemma 4.5 implies that $(-1)^d \psi^{(d)}$ also satisfies (3.2) with the same auxiliary function. Since

$$(-1)^k \psi^{(k)}(x) = \int_x^\infty (-1)^{k+1} \psi^{(k+1)}(t) dt$$

for all $k = 0, 1, \dots, d-1$, repeated use of Lemma 4.6 yields that ψ satisfies (3.2), again with the same auxiliary function. Hence $\psi \in \text{MDA}(\Lambda)$.

To prove sufficiency the procedure is simply reversed. In this direction the use of Lemma 4.4 is justified since $\psi^{(k)}$ are monotone by d -monotonicity for $k = 1, \dots, d-2$, and also for $k = d-1$ since $\psi^{(d-2)}$ is convex. For $k = d$, we use (4.5) together with the eventual monotonicity of \bar{F}' , which implies that $\psi^{(d)}$ also is eventually monotone—indeed, suppose $f(x) = x^\rho g(x)$ for some functions f and g on $(0, x^*)$, where $\rho \geq 0$. If for any $x_0 < x^*$ one can find $x, y \in (x_0, x^*)$, $x < y$, such that $g(x) < g(y)$, then f cannot be non-increasing on any left neighborhood of x^* , since clearly $f(x) < f(y)$. With $f = -\bar{F}'$ and $g = -\psi^{(d)}$, this gives a contradiction if $\psi^{(d)}$ is not eventually monotone. \square

4.3 Regular variation to the right, non-strict case

We start with a result that describes how non-strict d -monotone Archimedean generators and their derivatives behave in case they are regularly varying near the right endpoint.

Proposition 4.9 *Let $\psi = \mathfrak{W}_d F$ for some distribution function F with $F(0) = 0$, and suppose the right endpoint $x^* = \psi^{-1}(0)$ is finite.*

(i) If $\psi(x^* - x^{-1}) \in \text{RV}_{-\alpha}$, $\alpha > 0$, then

$$(-1)^k \psi^{(k)}(x^* - x^{-1}) \in \text{RV}_{-\alpha+k}$$

for $k = 1, \dots, d-1$. Moreover, in this case $\alpha \geq d-1$.

(ii) If $(-1)^{d-1} \psi^{(d-1)}(x^* - x^{-1}) \in \text{RV}_{-\alpha}$, $\alpha \geq 0$, then

$$(-1)^k \psi^{(k)}(x^* - x^{-1}) \in \text{RV}_{-\alpha-(d-1)+k}$$

for $k = 0, \dots, d-2$.

Proof. We first prove (i). Since $\psi^{(k)}$ is continuous for $k = 0, \dots, d-2$, $\psi^{(k)}(x^*) = 0$ for these k . Hence for $k = 1, \dots, d-1$,

$$\psi^{(k-1)}(x^* - x^{-1}) = - \int_x^\infty t^{-2} \psi^{(k)}(x^* - t^{-1}) dt. \quad (4.6)$$

Note that $\psi^{(d-1)}$ is in general only a density, not a derivative, of $\psi^{(d-2)}$ (see the comment after Definition 2.3). This causes no extra difficulties, since $\psi^{(d-2)}(x) = - \int_x^{x^*} g(s) ds$ implies (4.6) for $k = d-1$ upon changing variables to $t := (x^* - s)^{-1}$. We now proceed by induction over k . The induction assumption is that, as $x \rightarrow \infty$,

$$\psi^{(k-1)}(x^* - x^{-1}) \sim c_{k-1} x^{-\alpha+k-1} L(x), \quad -\alpha + k - 1 < 0 \text{ and } c_{k-1} \neq 0.$$

Here L is some slowly varying function. The Monotone Density Theorem and (4.6) then imply that $-x^{-2} \psi^{(k)}(x^* - x^{-1}) \sim (\alpha - k + 1) c_{k-1} x^{-\alpha+k-2} L(x)$, or, equivalently,

$$\psi^{(k)}(x^* - x^{-1}) \sim c_k x^{-\alpha+k} L(x), \quad (4.7)$$

where $c_k = (-\alpha + k - 1) c_{k-1} \neq 0$.

To complete the induction step, we must show that $-\alpha + k < 0$. So suppose first that $-\alpha + k > 0$. Then (4.7) implies that $(-1)^k \psi^{(k)}(x) \rightarrow +\infty$ as $x \rightarrow x^*$, which contradicts d -monotonicity. Hence $-\alpha + k \leq 0$ for $k = 1, \dots, d-1$, which in particular gives the last statement in part (i).

We still have to exclude $-\alpha + k = 0$. However, it is enough to show this for $k = 1, \dots, d-2$, since this still would allow us to reach (4.7) for $k = d-1$ and $c_{d-1} \neq 0$, but possibly with $-\alpha + d - 1 = 0$. So suppose $-\alpha + k = 0$ for any $k \in \{1, \dots, d-2\}$. Then $(-1)^k \psi^{(k)}(x^* - x^{-1}) \in \text{RV}_0$, but at the same time d -monotonicity implies that $(-1)^k \psi^{(k)}$ is convex for $k = 1, \dots, d-2$, and thus in particular

$$\begin{aligned} (-1)^k \psi^{(k)} \left(x^* - \frac{1}{2} x^{-1} \right) &= (-1)^k \psi^{(k)} \left(\frac{1}{2} (x^* - x^{-1}) + \frac{1}{2} x^* \right) \\ &\leq (-1)^k \frac{1}{2} \psi^{(k)}(x^* - x^{-1}) + (-1)^k \frac{1}{2} \psi^{(k)}(x^*) \\ &= (-1)^k \frac{1}{2} \psi^{(k)}(x^* - x^{-1}), \end{aligned}$$

since $\psi^{(k)}(x^*) = 0$. Equivalently,

$$2 \leq \frac{\psi^{(k)}(x^* - x^{-1})}{\psi^{(k)}(x^* - (2x)^{-1})}.$$

But here the right side tends to one as $x \rightarrow \infty$ since $(-1)^k \psi^{(k)}(x^* - x^{-1}) \in RV_0$, so the inequality is eventually violated. This gives the desired contradiction.

The statement that was to be shown now follows upon noting that the induction assumption is satisfied for $k = 1$ by hypothesis.

The proof of (ii) is simpler. This time the induction assumption is

$$(-1)^k \psi^{(k)}(x^* - x^{-1}) \in RV_{-\alpha - (d-1) + k},$$

where $k \leq d - 1$. It clearly holds for $k = d - 1$ by hypothesis. To prove the induction step, we use (4.6) together with the induction assumption to obtain

$$(-1)^{k-1} \psi^{(k-1)}(x^* - x^{-1}) = \int_x^\infty t^{-2} (-1)^k \psi^{(k)}(x^* - t^{-1}) dt = \int_x^\infty t^{-1-\alpha-d+k} L(t) dt$$

for some slowly varying L . Since $\alpha + d - k > 0$, Karamata's Theorem gives

$$(-1)^{k-1} \psi^{(k-1)}(x^* - x^{-1}) \sim \frac{1}{\alpha + d - k} x^{-\alpha - (d-1) + (k-1)} L(x),$$

and so $(-1)^{k-1} \psi^{(k-1)}(x^* - x^{-1}) \in RV_{-\alpha - (d-1) + (k-1)}$. \square

Note that for $\alpha \in (0, \infty)$, Proposition 4.9 says in particular that $\psi(x^* - x^{-1})$ is in $RV_{-\alpha-d+1}$ if and only if $\psi^{(d-1)}(x^* - x^{-1})$ is in $RV_{-\alpha}$. We use this fact to prove the following connection between the generator ψ and the radial part $R \sim F$. As in the case of rapid variation, Section 4.2, we restrict ourselves to differentiable distribution functions.

Proposition 4.10 *Let F be a distribution function with $F(0) = 0$ and right endpoint $x^* < \infty$, let $\psi = \mathfrak{W}_d F$ for $d \geq 2$ be its Williamson d -transform, and suppose that either F or $\psi^{(d-1)}$ is differentiable. Then, for $\alpha > 0$,*

$$\bar{F}(x^* - x^{-1}) \in RV_{-\alpha} \quad \text{if and only if} \quad \bar{\psi}(x^* - x^{-1}) \in RV_{-\alpha+d-1}.$$

Proof. We prove the “only if” part. By Lemma 4.7, both F and $\psi^{(d-1)}$ are differentiable, and hence

$$\bar{F}(x^* - x^{-1}) = - \int_x^\infty t^{-2} \bar{F}'(x^* - t^{-1}) dt,$$

since $\bar{F}(x^*) = 0$. By the Monotone Density Theorem, $-\bar{F}'(x^* - x^{-1}) \in RV_{-\alpha+1}$. Lemma 4.7 says that

$$-\bar{F}'(x^* - x^{-1}) = \frac{1}{(d-1)!} (x^* - x^{-1})^{d-1} (-1)^d \psi^{(d)}(x^* - x^{-1}),$$

which implies that also $(-1)^d \psi^{(d)}(x^* - x^{-1})$ is in $\text{RV}_{-\alpha+1}$, since $(x^* - x^{-1})^{d-1}$ tends to a constant and thus is slowly varying. Now,

$$(-1)^{d-1} \psi^{(d-1)}(x^* - x^{-1}) = \int_x^\infty t^{-2} (-1)^d \psi^{(d)}(x^* - t^{-1}) dt,$$

where the right side is equal to $\int_x^\infty t^{-\alpha-1} L(t) dt$ for some slowly varying L , due the regular variation of $(-1)^d \psi^{(d)}(x^* - x^{-1})$. Karamata's Theorem then implies that $(-1)^{d-1} \psi^{(d-1)}(x^* - x^{-1}) \in \text{RV}_{-\alpha}$, and we may apply Proposition 4.9 to reach the desired conclusion.

To prove the converse, one argues analogously, but in the reverse order. First one establishes regular variation of $\psi^{(d-1)}(x^* - x^{-1})$ using Proposition 4.9. By the differentiability assumption, one can differentiate once more and use the Monotone Density Theorem together with Lemma 4.7 to obtain regular variation of $-\bar{F}'(x^* - x^{-1})$. Karamata's theorem then gives the conclusion about \bar{F} . \square

The intuition behind this result is that when F (and hence ψ) has a finite right endpoint x^* , then the asymptotic behavior of F near x^* is the same as the asymptotic behavior of $\psi^{(d-1)}$ near x^* . This is different from the case when x^* is infinite—then F and ψ have the same asymptotics, according to Proposition 4.3.

4.4 Behavior at the origin

We conclude the investigation of the asymptotics of the Williamson d -transform by considering the properties near the origin.

Proposition 4.11 *Let R be a random variable on $(0, \infty)$ with Williamson d -transform ψ . Then*

- (i) *If $1/R \in \text{RV}_{-\alpha}$, $\alpha \in (0, 1)$, then $1 - \psi(x^{-1}) \in \text{RV}_{-\alpha}$.*
- (ii) *If $1/R \in \text{RV}_{-\alpha}$, $\alpha \in [1, \infty]$, then $1 - \psi(x^{-1}) \in \text{RV}_{-1}$.*
- (iii) *If $1 - \psi(x^{-1}) \in \text{RV}_{-\alpha}$, $\alpha \in (0, 1)$, then $1/R \in \text{RV}_{-\alpha}$.*

Proof. Let S_1 be a $\text{Beta}(1, d-1)$ distributed random variable independent of R . Since S_1 is continuous we have

$$\mathbb{P}[1/S_1 > x] = \mathbb{P}[S_1 \leq 1/x] = 1 - (1 - x^{-1})^{d-1} \sim x^{-1}, \quad x \rightarrow \infty,$$

which implies that $1/S_1 \in \text{RV}_{-1}$. Furthermore, it follows from the discussion after Proposition 2.5 that $1 - \psi(x^{-1})$ is the survival function of $1/(RS_1)$.

To show (i), we pick an $\varepsilon > 0$ such that $\alpha + \varepsilon < 1$, which is possible since $\alpha < 1$. Then because $1/S_1 \in \text{RV}_{-1}$, $\mathbb{E}[(1/S_1)^{\alpha+\varepsilon}] < \infty$. Now, $1/R \in \text{RV}_{-\alpha}$ by assumption, so Breiman's result, Lemma 4.1, implies that $1/(RS_1) \in \text{RV}_{-\alpha}$, or, in other words, $1 - \psi(x^{-1}) \in \text{RV}_{-\alpha}$.

As for (ii), consider first the case when $\alpha \in (1, \infty]$. Then because $1/R \in \text{RV}_{-\alpha}$, $\mathbb{E}[(1/R)^{1+\varepsilon}] < \infty$ whenever ε is such that $1 + \varepsilon < \alpha$. One may then use Breiman's result again to conclude that $1/(RS_1) \in \text{RV}_{-1}$, i.e. $1 - \psi(x^{-1}) \in \text{RV}_{-1}$. If $\alpha = 1$, Breiman's result is not applicable. Instead we show explicitly that in this case,

$$L(x) := \frac{1 - \psi(1/x)}{1/x}, \quad x \in (0, \infty),$$

is slowly varying. To do so, first use the definition of the Williamson d -transform and the binomial expansion theorem to write

$$L(x) = - \sum_{k=1}^{d-1} \binom{d-1}{k} (-1)^k x^{-k+1} \mathbb{E}[R^{-k} \mathbb{1}_{\{R > 1/x\}}], \quad x \in (0, \infty).$$

Further, denote by \bar{G} the survival function of $1/R$ and integrate by parts to obtain

$$\mathbb{E}[R^{-k} \mathbb{1}_{\{R > 1/x\}}] = - \int_0^x y^k d\bar{G}(y) = -x^k \bar{G}(x) + k \int_0^x \bar{G}(y) y^{k-1} dy.$$

Thus $L(x)$ may be expressed as $L(x) = A(x)\{B_1(x) + \cdots + B_{d-1}(x)\}$ where

$$A(x) = \int_0^x \bar{G}(y) dy, \quad B_1(x) = (d-1) \left\{ 1 - \frac{x\bar{G}(x)}{A(x)} \right\}$$

and

$$B_k(x) = \binom{d-1}{k} (-1)^{k-1} \frac{x\bar{G}(x)}{A(x)} \left(-1 + \frac{k \int_0^x \bar{G}(y) y^{k-1} dy}{x^k \bar{G}(x)} \right), \quad k = 2, \dots, (d-1).$$

Karamata's Theorem now implies that

$$\lim_{x \rightarrow \infty} \frac{x\bar{G}(x)}{A(x)} = 0, \quad \lim_{x \rightarrow \infty} \frac{k \int_0^x \bar{G}(y) y^{k-1} dy}{x^k \bar{G}(x)} = \frac{k}{k-1}, \quad k = 2, \dots, (d-1).$$

Thus as $x \rightarrow \infty$, $B_1(x) \rightarrow d-1$ while $B_k(x) \rightarrow 0$ for $k = 2, \dots, (d-1)$. Since $A(x)$ is slowly varying by Karamata's Theorem, $L(x)$ is a product of two slowly varying functions and therefore itself slowly varying.

Part (iii) is based on the Monotone Density Theorem. Suppose $1 - \psi(x^{-1}) = x^{-\alpha} L(x)$, with L slowly varying. Since $\psi(0) = 1$ we get

$$1 - \psi(x^{-1}) = - \int_0^{1/x} \psi'(y) dy = - \int_x^\infty y^{-2} \psi'(y^{-1}) dy.$$

The Monotone Density Theorem gives $-x^{-1} \psi'(x^{-1}) \sim \alpha x^{-\alpha} L(x)$ as $x \rightarrow \infty$. Suppose now that $(-1)^k x^{-k} \psi^{(k)}(x^{-1}) \sim c_k x^{-\alpha} L(x)$ for some $1 \leq k \leq d-2$, where c_k is a constant. Since $\psi^{(k)}(x) = - \int_x^\infty \psi^{(k+1)}(y) dy$ we get, after changing variables,

$$(-1)^k \psi^{(k)}(x^{-1}) = \int_0^x (-1)^{k+1} \psi^{(k+1)}(y^{-1}) y^{-2} dy.$$

By the Monotone Density Theorem, $(-1)^{k+1}\psi^{(k+1)}(x^{-1})x^{-2} \sim c_k(-\alpha+k)x^{-\alpha+k-1}L(x)$, and so by induction we obtain, for $k = 0, 1, \dots, d-1$,

$$(-1)^k x^{-k} \psi^{(k)}(x^{-1}) \sim c_k x^\alpha L(x), \quad x \rightarrow \infty,$$

where the c_k satisfy $c_0 = 1$, $c_1 = \alpha$ and $c_{k+1} = (k - \alpha)c_k > 0$. The inversion formula for the Williamson d -transform now yields

$$\begin{aligned} F(x^{-1}) &= 1 - \psi(x^{-1}) - \sum_{k=1}^{d-1} \frac{(-1)^k x^{-k} \psi^{(k)}(x^{-1})}{k!} \\ &\sim \sum_{k=0}^{d-1} \frac{1}{k!} c_k x^\alpha L(x) = B(1 - \psi(x^{-1})), \quad x \rightarrow \infty, \end{aligned}$$

where the constant B is positive since each c_k is. \square

Suppose that ψ is a d -monotone Archimedean generator. In particular, this implies that $\tilde{\psi}(x) := \psi(-x)$ is a distribution function on $(-\infty, 0]$. Our final result shows that d -monotonicity of ψ restricts the possible extremal behavior of $\tilde{\psi}$. In particular, it shows that Proposition 4.11 indeed covers (almost) all relevant cases (the exception being $\alpha = 1$ in part (iii), which is not covered).

Proposition 4.12 *Let ψ be a d -monotone Archimedean generator and $\tilde{\psi}$ as above. Then*

- (i) *If $1 - \psi(x^{-1}) \in \text{RV}_{-\alpha}$, then $\alpha \in [0, 1]$.*
- (ii) *If the extremal behavior of $\tilde{\psi}$ is non-degenerate, then $\tilde{\psi}$ is in the maximum domain of attraction of the Weibull distribution with the parameter range restricted to $(0, 1]$.*

Proof. Part (i) stems from the fact that if $\alpha > 1$,

$$\lim_{x \downarrow 0} \frac{1 - \psi(x/2)}{1 - \psi(x)} = 2^{-\alpha} < \frac{1}{2}.$$

In particular, therefore, there exists an $x_0 \in (0, \infty)$ such that $\psi(x/2) > \{\psi(0) + \psi(x)\}/2$ whenever $x \in (0, x_0)$. But ψ is d -monotone and hence convex, so we obtain a contradiction.

Turning to (ii), observe that since the upper endpoint of $\tilde{\psi}$ is finite, $\tilde{\psi}$ cannot be in the Fréchet domain of attraction. Furthermore, a necessary condition for $\tilde{\psi} \in \text{MDA}(\Lambda)$ is that $1 - \psi(x^{-1}) \in \text{RV}_{-\infty}$, which is excluded by (i). So $\tilde{\psi}$ is in the Weibull domain. This implies in particular that $1 - \psi(x^{-1}) \in \text{RV}_{-\alpha}$ for some $\alpha \in (0, \infty)$. Due to (i) however, $\alpha > 1$ is not possible. \square

4.5 Conditions on the inverse generator

The most common approach found in the literature for dealing with Archimedean copulas is to express them in terms of the inverse generator $\phi = \psi^{-1}$. We have already seen some of the virtues of working with the generator ψ instead, and the advantages will become even more apparent later on. However, in order to remain compatible with existing conventions we here give equivalent conditions on ϕ for the results in Sections 4.1–4.4.

Proposition 4.13 *Let ψ be an Archimedean generator, and let $\phi = \psi^{-1}$ be the inverse generator.*

(i) $\psi \in \text{RV}_{-\alpha}$, $\alpha \in (0, \infty)$ if and only if $\phi(1/s) \in \text{RV}_{1/\alpha}$, or, equivalently,

$$\lim_{s \downarrow 0} \frac{\phi(cs)}{\phi(s)} = c^{-1/\alpha} \quad \text{for all } c > 0.$$

(ii) $\psi(x^* - x^{-1}) \in \text{RV}_{-\alpha}$, $\alpha \in [d-1, \infty)$, if and only if $(x^* - \phi(1/s))^{-1} \in \text{RV}_{1/\alpha}$, or, equivalently,

$$\lim_{s \downarrow 0} \frac{x^* - \phi(cs)}{x^* - \phi(s)} = c^{1/\alpha} \quad \text{for all } c > 0.$$

Here x^* is the right endpoint of ψ .

(iii) $1 - \psi(x^{-1}) \in \text{RV}_{-\alpha}$, $\alpha \in (0, 1]$, if and only if $\phi(1 - s^{-1}) \in \text{RV}_{-1/\alpha}$, or, equivalently,

$$\lim_{s \downarrow 0} \frac{\phi(1 - cs)}{\phi(1 - s)} = c^{1/\alpha} \quad \text{for all } c > 0.$$

(iv) $\psi \in \text{MDA}(\Lambda)$ if and only if $\phi(1/s)$ is in de Haan's Π -class¹. The auxiliary function $a(x)$ of $\psi(x)$ and $c(s)$ of $\phi(1/s)$ are related through $c(s) = a(\phi(1/s))$ and $a(x) = c(1/\psi(x))$.

To simplify matters, we need the following adaptation of a theorem on inverses of regularly varying functions:

Lemma 4.14 *Suppose f is invertible and $\alpha \in (0, \infty)$. Then $f \in \text{RV}_{-\alpha}$, $\alpha > 0$ if and only if $f^{-1}(s^{-1}) \in \text{RV}_{1/\alpha}$, or, equivalently,*

$$\lim_{s \downarrow 0} \frac{f^{-1}(cs)}{f^{-1}(s)} = c^{-1/\alpha}, \quad \text{for all } c > 0.$$

¹A non-negative, non-decreasing function g is in de Haan's Π -class if $\lim_{s \rightarrow \infty} \{g(ts) - b(s)\}/c(s) = \log t$, $t \in \mathbb{R}$, for some functions $b(s) \in \mathbb{R}$ and $c(s) > 0$. The function $c(s)$ is called the *auxiliary function* of g . See Resnick [25], Section 0.4.3 for an concise account of the relevant theory.

Proof. $f \in \text{RV}_{-\alpha}$ if and only if $1/f \in \text{RV}_{\alpha}$. An easy computation shows that $[1/f]^{-1}(s) = f^{-1}(s^{-1})$, which by Theorem 1.5.12 in Bingham et al. [2] is equivalent to $f^{-1}(s^{-1}) \in \text{RV}_{1/\alpha}$. The last equivalent formulation follows directly from the definition of regular variation upon changing the variable from s to s^{-1} . \square

Proof. (of Proposition 4.13) Part (i) follows directly from Lemma 4.14 since ψ is strict and hence invertible. For Part (ii), note that $x^* - x^{-1} \in \{\psi > 0\}$. On this set ψ is invertible, and the inverse of $\psi(x^* - x^{-1})$ is $1/(x^* - \phi(s^{-1}))$. Now apply Lemma 4.14. Part (iii) follows in the same fashion since $\psi(x^{-1})$ is invertible for x sufficiently large, with inverse $\phi(1 - s^{-1})$. For Part (iv), Proposition 0.10 in Resnick [25] says that $\psi \in \text{MDA}(\Lambda)$ if and only if $(1/\psi)^{\leftarrow}(s) = \inf\{x \geq 0 \mid 1/\psi(x) \geq s\}$ is in de Haan's Π -class. But since ψ is continuous, we have $\inf\{x \geq 0 \mid 1/\psi(x) \geq s\} = \inf\{x \geq 0 \mid 1/\psi(x) = s\}$, which is equal to $\phi(s^{-1})$ by definition. The relationship between the auxiliary functions follows from Proposition 0.9 in [25], where, in order to obtain the expression for $a(x)$, we used that $[\phi(1/s)]^{\leftarrow} = 1/\psi$ since ϕ is a right inverse of ψ : $\phi(1/s) = x$ implies $1/s = \psi(x)$. \square

Chapter 5

Coefficients of tail dependence

5.1 Coefficient of the first kind

An important problem in probability theory and statistics is to find simple measures, preferably scalar, that quantify the strength of dependence between random variables. In the context of extremes, one may ask about the likelihood of observing large values in one component X_1 given that another component X_2 is large. One such measure is the *coefficient of (upper) tail dependence* λ_u , defined as

$$\lambda_u := \lim_{q \rightarrow 1} \mathbb{P}[X_1 > F_1^{\leftarrow}(q) \mid X_2 > F_2^{\leftarrow}(q)],$$

where F_1 and F_2 are the marginal distribution functions of X_1 and X_2 , respectively, and the limit is assumed to exist (this is not always the case!). The *coefficient of lower tail dependence*, λ_l , is defined analogously:

$$\lambda_l := \lim_{q \rightarrow 0} \mathbb{P}[X_1 \leq F_1^{\leftarrow}(q) \mid X_2 \leq F_2^{\leftarrow}(q)].$$

Note that dependence is measured on a quantile scale. It is therefore not surprising that these measures can be expressed in terms of the associated copula C . Indeed, using the definition of the generalized inverse, one easily checks the validity of

$$\mathbb{P}[X_1 \leq F_1^{\leftarrow}(q) \mid X_2 \leq F_2^{\leftarrow}(q)] = \frac{\mathbb{P}[F_1(X_1) \leq q, F_2(X_2) \leq q]}{\mathbb{P}[F_2(X_2) \leq q]} = \frac{C(q, q)}{q}.$$

Similarly one can show that λ_u also depends only on the copula. In the case where C is Archimedean, λ_u and λ_l have simple forms:

Proposition 5.1 *Provided the limits exist, the coefficients of upper and lower tail de-*

pendence of an Archimedean copula with generator ψ are given by

$$\begin{aligned}\lambda_u &= 2 - \lim_{t \rightarrow 0} \frac{1 - \psi(2t)}{1 - \psi(t)} \\ \lambda_l &= \begin{cases} \lim_{t \rightarrow \infty} \frac{\psi(2t)}{\psi(t)}, & \text{if } \psi \text{ is strict} \\ 0, & \text{otherwise} \end{cases}\end{aligned}$$

Proof. Let (U_1, \dots, U_d) be distributed according to C . Due to exchangeability, all two-dimensional marginals are given by $C(u_1, u_2, 1, \dots, 1) = \psi(\psi^{-1}(u_1) + \psi^{-1}(u_2))$. The expression for λ_u and the fact that the marginals are uniform yield

$$\begin{aligned}\lambda_u &= \lim_{q \rightarrow 1} \frac{\mathbb{P}[U_1 > q, U_2 > q]}{\mathbb{P}[U_2 > q]} \\ &= \lim_{q \rightarrow 1} \frac{1 - \mathbb{P}[U_1 \leq q] - \mathbb{P}[U_2 \leq q] + \mathbb{P}[U_1 \leq q, U_2 \leq q]}{1 - \mathbb{P}[U_2 \leq q]} \\ &= 2 - \lim_{t \rightarrow 0} \frac{1 - \psi(2t)}{1 - \psi(t)}.\end{aligned}$$

Here we used that $q = \psi \circ \psi^{-1}(q)$ and that $t := \psi^{-1}(q) \rightarrow 0$ as $q \rightarrow 1$. Similarly,

$$\lambda_l = \lim_{q \rightarrow 0} \frac{\mathbb{P}[U_1 \leq q, U_2 \leq q]}{\mathbb{P}[U_2 \leq q]} = \lim_{q \rightarrow 0} \frac{\psi(2\psi^{-1}(q))}{q} = \lim_{t \rightarrow \psi^{-1}(0)} \frac{\psi(2t)}{\psi(t)}.$$

If ψ is strict, then $\psi^{-1}(0) = \infty$ and we get the result. If ψ is non-strict, then $c := \psi^{-1}(0) < \infty$. So for all $t \in [c/2, c)$, the numerator $\psi(2t)$ is zero, while the denominator is non-zero. Hence the limit must be zero. \square

Note that the generator in Example 4.2 provides an example where the index of lower tail dependence does not exist. However, the results from Section 4 allow us to give sufficient conditions for the existence of the coefficients of tail dependence and compute their values in terms of the corresponding radial part R .

Corollary 5.2 *Suppose the Archimedean copula C has generator ψ , and let $R \sim F = \mathfrak{W}_d^{-1}\psi$ be the corresponding radial distribution.*

(i) *If $1/R \in \text{RV}_{-\alpha}$, $\alpha \in (0, \infty)$, then $\lambda_u = 2 - 2^\beta$, where $\beta = \alpha \wedge 1$.*

(ii) *If $R \in \text{RV}_{-\alpha}$, $\alpha \in [0, \infty)$, then $\lambda_l = 2^{-\alpha}$.*

Proof. For (i), we observe that $1 - \psi(x^{-1}) \in \text{RV}_{-\beta}$ according to Proposition 4.11, and hence

$$\lim_{t \rightarrow 0} \frac{1 - \psi(2t)}{1 - \psi(t)} = \lim_{x \rightarrow \infty} \frac{1 - \psi((x/2)^{-1})}{1 - \psi(x^{-1})} = 2^\beta,$$

by definition of regular variation. The expression in Proposition 5.1 for λ_u gives the result. For (ii), the result is immediate from Proposition 4.3, the definition of regular variation, and Proposition 5.1. \square

We remark that this result does not give conditions on R ensuring that $\lambda_u = 1$. This is due to the limitation $\alpha > 0$ in Proposition 4.11 (i).

If $\psi \in \text{RV}_{-\infty}$, then Proposition 5.1 clearly implies that $\lambda_l = 0$. Proposition 4.8 gives one sufficient condition for this to happen, namely that the corresponding radial part R is in the Gumbel domain of attraction and has differentiable distribution function with eventually monotone derivative.

Note that the converse of Corollary 5.2 does not hold—knowing e.g. λ_l only gives information about the limit of $\psi(cx)/\psi(x)$ for $c = 2$, which is not enough to deduce regular variation of ψ (and thus of R). The situation is analogous for λ_u .

A natural extension of the coefficients of tail dependence for high-dimensional distributions is

$$\lambda_l^{m,n} := \lim_{q \rightarrow 0} \mathbb{P}[X_1 \leq F_1^{\leftarrow}(q), \dots, X_m \leq F_m^{\leftarrow}(q) \mid X_{m+1} \leq F_{m+1}^{\leftarrow}(q), \dots, X_{m+n} \leq F_{m+n}^{\leftarrow}(q)],$$

and analogously for $\lambda_u^{m,n}$. As in Proposition 5.1 one shows that for Archimedean copulas, we have the expression

$$\lambda_l^{m,n} = \lim_{t \rightarrow \infty} \frac{\psi((m+n)t)}{\psi(nt)}$$

if ψ is strict. If ψ is regularly varying with index $-\alpha$, $\alpha \in [0, \infty)$ (which happens if R has the same property), we conclude as before that

$$\lambda_l^{m,n} = \left(\frac{m}{n} + 1\right)^{-\alpha}. \quad (5.1)$$

The expression for λ_u is possible to derive, but has a more complicated form since its derivation requires that survival functions of the copula C be expressed in terms of marginal distribution functions of different dimension. This can be done using the so-called sieve formula, but the calculations are cumbersome.

As a final observation we note that although there is no hope of a finding a converse to Corollary 5.2, the situation is in a sense different for $\lambda_l^{m,n}$. For, if we know that $\lambda_l^{m,n}$ has the form (5.1) for all $m, n \in \mathbb{N}$, then ψ must be regularly varying with index $-\alpha$. This follows from Theorem 1.4.3 in Bingham et al. [2], which says that $\psi \in \text{RV}_{-\alpha}$ if and only if $\psi(cx)/\psi(x) \rightarrow c^{-\alpha}$ for all c in a dense subset of $(0, \infty)$, provided the condition

$$\limsup_{c \downarrow 1} \psi^*(c) \leq 1, \quad \text{where } \psi^*(c) := \limsup_{x \rightarrow \infty} \psi(cx)/\psi(x),$$

is satisfied. But this automatically holds since ψ is decreasing and hence $\psi^*(c) \leq 1$ for all $c \geq 1$. Furthermore, although (5.1) only gives $\psi(cx)/\psi(x) \rightarrow c^{-\alpha}$ for all rational c in $(1, \infty)$, this is enough since for $0 < c < 1$ rational, it implies

$$\lim_{x \rightarrow \infty} \psi(cx)/\psi(x) = \lim_{x \rightarrow \infty} \psi(x)/\psi(x/c) = \left\{ \lim_{x \rightarrow \infty} \psi(x/c)/\psi(x) \right\}^{-1} = c^{-\alpha}.$$

Of course, this way of inferring properties of ψ , and hence R , based on the behavior of $\lambda_l^{m,n}$ is of little practical importance since *all* these coefficients must satisfy (5.1). The corresponding estimation problem is obviously infeasible. Moreover, the interpretation of $\lambda_l^{m,n}$ as describing multivariate marginal tail behavior only remains valid if ψ generates a copula in *any* dimension, and thus is completely monotone.

5.2 Coefficient of the second kind

When a coefficient of tail dependence is zero, this is interpreted as *asymptotic independence*¹ of the tails of the marginals, in the sense that unlikely events tend to occur more or less independently of each other. The “more or less” part is of course not completely satisfactory, and supplementary measures of tail dependence have been suggested. One such measure was introduced by Coles et al. [7], and following Demarta [8] we call it the (*upper*) *coefficient of tail dependence of the second kind*. It is defined as

$$\bar{\lambda}_u = \lim_{q \rightarrow 0} \frac{\log(\mathbb{P}[X_1 > F_1^{\leftarrow}(q)]\mathbb{P}[X_2 > F_2^{\leftarrow}(q)])}{\log(\mathbb{P}[X_1 > F_1^{\leftarrow}(q), X_2 > F_2^{\leftarrow}(q)])} - 1,$$

with the lower coefficient $\bar{\lambda}_l$ being defined analogously. These quantities compare, on a logarithmic scale, the joint exceedance probabilities under the actual dependence structure with the exceedance probabilities one would observe if the components were independent. Expressed in terms of the copula C of (X_1, X_2) , the tail dependence coefficients of the second kind are

$$\bar{\lambda}_u = \lim_{q \rightarrow 1} \frac{2 \log(1 - q)}{\log(1 - 2q + C(q, q))} - 1 \quad \text{and} \quad \bar{\lambda}_l = \lim_{q \rightarrow 0} \frac{2 \log q}{\log C(q, q)} - 1. \quad (5.2)$$

We first consider the lower coefficient $\bar{\lambda}_l$. It turns out that when $\psi \in \text{RV}_{-\alpha}$ for $\alpha > 0$, and hence $\lambda_l > 0$, $\bar{\lambda}_l$ gives no additional information:

Proposition 5.3 *Suppose $\psi \in \text{RV}_{-\alpha}$, $\alpha > 0$. Then $\bar{\lambda}_l = 1$.*

Proof. Change the variable to $x := \psi^{-1}(q)$ in (5.2) and apply L'Hôpital's rule:

$$\lim_{q \rightarrow 0} \frac{2 \log q}{\log C(q, q)} = \lim_{x \rightarrow \infty} \frac{2 \log \psi(x)}{\log \psi(2x)} = \lim_{x \rightarrow \infty} \frac{\psi(2x) \psi'(2x)}{\psi(x) \psi'(x)}.$$

Now, ψ' is monotone by d -monotonicity, and $\psi(x) = -\int_x^\infty \psi'(t)dt$. Since $\alpha \neq 0$, the Monotone Density Theorem implies that $-\psi' \in \text{RV}_{-\alpha-1}$, and so the limit above is equal to $2^{-\alpha}/2^{-\alpha-1} = 2$. By (5.2), then, $\bar{\lambda}_l = 1$. \square

In fact one can easily show that $\bar{\lambda}_l$ is always one whenever $\lambda_l > 0$ (see Demarta [8], Lemma 2.1 (iii) for the corresponding property in the case of $\bar{\lambda}_u$). If C is Archimedean

¹This term has several different meanings depending on the context. Here we use it a vague sense to support the intuition about the coefficients of tail dependence.

with strict generator ψ , and ψ is in the Gumbel domain of attraction (and hence in $\text{RV}_{-\infty}$, implying that $\lambda_l = 0$), we can be more precise about the behavior of $\bar{\lambda}_l$.

Proposition 5.4 *Suppose C is a bivariate Archimedean copula whose generator ψ is strict and in the Gumbel domain of attraction with auxiliary function a . If $a \in \text{RV}_\beta$, then*

$$\bar{\lambda}_l = 2^{-\beta} - 1.$$

Proof. As in the preceding proof we obtain

$$\lim_{q \rightarrow 0} \frac{2 \log q}{\log C(q, q)} = \lim_{x \rightarrow \infty} \frac{\psi'(x)/\psi(x)}{\psi'(2x)/\psi(2x)}.$$

By Theorem 3.5 and the subsequent remark, we have $a(x) \sim -\psi(x)/\psi'(x)$, whence the result follows due to the assumption on a and the definition of regular variation. \square

These results are in line with the intention that $\bar{\lambda}_l$ should allow to differentiate between different degrees of asymptotic independence (in the above vague sense). The conditions on ψ in Proposition 5.4 can immediately be translated into conditions on the radial part R using Proposition 4.8.

As mentioned above, there is a counterpart to Proposition 5.3 for the upper coefficient $\bar{\lambda}_u$. More precisely, $\bar{\lambda}_u = 1$ as soon as $\lambda_u > 0$. A sufficient condition for this to happen is, according to Corollary 5.2, that the radial part R corresponding to ψ satisfies $1/R \in \text{RV}_{-\alpha}$ with $\alpha \in (0, 1)$. We have not been able to find a counterpart to Proposition 5.4, i.e. to deduce what happens in the case $\alpha \in [1, \infty)$, and hence $\lambda_u = 0$.

We conclude this chapter with a short discussion on the relationship between distributions, their copulas, survival copulas, and coefficients of tail dependence. The tail dependence coefficients are sometimes defined as a quantity relating to the *distribution* F of a random vector, but having the property that it only depends on the (ordinary or survival) copula of that distribution. That is how the concept was approached here. However, many authors define the tail dependence coefficients as quantities relating directly to a particular *copula* C , regardless of whether this copula in a specific modeling context appears as the (ordinary) copula of a distribution under consideration, or as its survival copula. This latter convention has the consequence that if λ_u , say, is the upper tail dependence coefficient of C , and C is the (ordinary) copula of F , then λ_u describes the *upper* tail of F . If, on the other hand, C is the survival copula of F , then λ_u describes the *lower* tail of F . When dealing with Archimedean copulas, this can become confusing since these copulas appear as the *survival copulas* of ℓ_1 -norm symmetric distributions. Hence the quantities λ_l and $\bar{\lambda}_l$ treated in this chapter both describe the *upper* tail of the corresponding ℓ_1 -norm symmetric distribution, while λ_u and $\bar{\lambda}_u$ describe its *lower* tail.

Chapter 6

Extreme value limits of Archimedean copulas

A d -dimensional copula C with the property that $C^n(\mathbf{u}^{1/n}) = C(\mathbf{u})$ for all $n > 0$ and $\mathbf{u} \in [0, 1]^d$ is called *extreme value* (EV) copula. These copulas arise when one considers maxima (or, as we will see, minima) of multivariate distributions. It is well-known (see Genest and Rivest [15]) that if C is Archimedean, it belongs to the Gumbel family, i.e. its generator is $\psi_\theta^{\text{Gu}}(x) = \exp(-x^{1/\theta})$ for some $\theta \geq 1$. The language of ℓ_1 -norm symmetric distributions makes it simple to understand why this is to be expected. This will be apparent in Section 6.1, where we consider the limiting distributions of componentwise minima of ℓ_1 -norm symmetric distributions. In Section 6.2 we consider the maxima of these distributions. Here the limit turns out to be the so-called Galambos copula (see [14] and [5]).

6.1 Componentwise minima of ℓ_1 -norm symmetric distributions

Consider an i.i.d. sequence $\mathbf{X}_1, \mathbf{X}_2, \dots$ of \mathbb{R}^d -valued random vectors. The corresponding *componentwise n -block minima* are defined as

$$\mathbf{W}_n = (W_{1,n}, \dots, W_{d,n}) := (\min\{X_{1,k}\}_{k=1}^n, \dots, \min\{X_{d,k}\}_{k=1}^n).$$

In the sequel it will always be assumed that each $\mathbf{X}_n = (X_{1,n}, \dots, X_{d,n})$ follows a d -dimensional ℓ_1 -norm symmetric distribution with radial part $R \sim F$, where $R > 0$ a.s., and joint survival function \bar{H} . By Theorem 2.12 it then follows that $\bar{H}(\mathbf{x}) = \psi(\|\mathbf{x}\|_1)$ for $\mathbf{x} \in \mathbb{R}_+^d$, where $\psi = \mathfrak{W}_d F$. Moreover, ψ is then the generator of the survival copula C of \mathbf{X}_n , which is Archimedean. We now fix this notation and do not repeat it at each instance it appears in the sequel.

The ℓ_1 -norm symmetric distributions have the following closure property under the operation of forming minima:

Lemma 6.1 *The n -block minima \mathbf{W}_n are again ℓ_1 -norm symmetric. The corresponding survival copula generators are ψ^n .*

Proof. For each n , \mathbf{W}_n is clearly confined to the interior of \mathbb{R}_+^d with probability one. Its survival function is

$$\begin{aligned}\mathbb{P}[\mathbf{W}_n > \mathbf{x}] &= \mathbb{P}[\{\min\{X_{1,k}\}_{k=1}^n > x_1\} \cap \cdots \cap \{\min\{X_{d,k}\}_{k=1}^n > x_d\}] \\ &= \mathbb{P}\left[\bigcap_{i=1,\dots,d} \bigcap_{k=1,\dots,n} \{X_{i,k} > x_i\}\right] \\ &= \left(\mathbb{P}\left[\bigcap_{i=1,\dots,d} \{X_{i,1} > x_i\}\right]\right)^n \\ &= \psi^n(\|\mathbf{x}\|_1),\end{aligned}$$

where the third equality uses mutual independence of $\mathbf{X}_1, \dots, \mathbf{X}_n$. The result now follows by Theorem 2.12. \square

Lemma 6.2 *Suppose $1 - \psi(x^{-1}) \in RV_{-1/\theta}$ for some $\theta < \infty$. Then there exists a sequence (c_n) of positive constants such that $c_n^{-1}W_{1,n} \Rightarrow Y$ as $n \rightarrow \infty$, where the survival function of Y is ψ_θ^{Gu} .*

Proof. Using the identity $\max(-x) = -\min(x)$ we get $-W_{1,n} = \max\{-X_{1,k}\}_{k=1}^n$. The random variables $-X_{1,k}$ are i.i.d. with distribution function $\psi(-x)$ for $x \leq 0$, which has right endpoint zero. Moreover, since $1 - \psi(x^{-1}) \in RV_{-1/\theta}$ by assumption, $\psi(-x)$ is in the Weibull domain of attraction by Theorem 3.5. By Proposition 3.6, there exist positive constants c_n such that $-c_n^{-1}W_{1,n} \Rightarrow -Y \sim \exp(-(-x)^{1/\theta}) = \psi_\theta^{\text{Gu}}(-x)$. The result now follows upon noting that, for $x \geq 0$,

$$\mathbb{P}[Y > x] = \mathbb{P}[-Y \leq -x] = \psi_\theta^{\text{Gu}}(-(-x)),$$

where we used the continuity of ψ_θ^{Gu} . \square

Before continuing we note that Proposition 4.12 implies that $1/\theta \in [0, 1]$. Since the hypothesis of Lemma 6.2 excludes $\theta = \infty$, we always have $\theta \in [1, \infty)$, which is consistent with the requirements of the Gumbel generator.

We are now in a position to prove the next result on the limiting distribution of normalized n -block minima of ℓ_1 -norm symmetric distributions. As a corollary we obtain the extreme value limits of Archimedean copulas.

Proposition 6.3 *Suppose $1 - \psi(x^{-1}) \in RV_{-1/\theta}$ for some $\theta < \infty$. Then there exists a sequence (c_n) of positive constants such that $c_n^{-1}\mathbf{W}_n \Rightarrow \mathbf{W}_0$ as $n \rightarrow \infty$, where \mathbf{W}_0 is ℓ_1 -norm symmetric with survival copula generator ψ_θ^{Gu} .*

Proof. For any constant $a > 0$, Lemma 6.1 implies that the survival function of $a^{-1}\mathbf{W}_n$ is $\psi^n(a\|\mathbf{x}\|_1)$, with the one-dimensional marginals thus being $\psi^n(ax)$. With c_n as in Lemma 6.2, we therefore conclude that the functions $x \mapsto \psi^n(c_n x)$ converge weakly to ψ_θ^{Gu} as $n \rightarrow \infty$. Hence $\mathbb{P}[c_n^{-1}\mathbf{W}_n > \mathbf{x}] = \psi^n(c_n\|\mathbf{x}\|_1) \rightarrow \psi_\theta^{\text{Gu}}(\|\mathbf{x}\|_1)$ weakly as $n \rightarrow \infty$, which means precisely that $c_n^{-1}\mathbf{W}_n \Rightarrow \mathbf{W}_0$. Theorem 2.12 implies that \mathbf{W}_0 has the desired properties. \square

Corollary 6.4 *Suppose C is an Archimedean copula whose generator ψ satisfies $1 - \psi(x^{-1}) \in \text{RV}_{-1/\theta}$ for some $\theta < \infty$. Then $C^n(\mathbf{u}^{1/n}) \rightarrow C_\theta^{\text{Gu}}(\mathbf{u})$ pointwise.*

Proof. Let the \mathbf{X}_n above follow the ℓ_1 -norm symmetric distribution whose survival copula generator is ψ . By Lemma 6.1, the survival copula of \mathbf{W}_n , and hence of $a\mathbf{W}_n$ for any $a > 0$, is

$$\psi^n(\psi^{-1}(u_1^{1/n}) + \dots + \psi^{-1}(u_1^{1/n})) = C^n(\mathbf{u}^{1/n}).$$

Weak convergence of a sequence of joint distributions toward a limiting distribution implies pointwise convergence of the survival copulas toward the survival copula of the limiting distribution (see e.g. Feidt et al. [12]). By Proposition 6.3, the limiting distribution has survival copula C_θ^{Gu} , with θ as in the statement of the corollary. The result follows. \square

In the case of ℓ_1 -norm symmetric distributions, it is the *survival copula* that is Archimedean. This is the reason why we considered the *minima* of these distributions in order to arrive at the conclusion in Corollary 6.4. However, in a situation where an Archimedean copula C is the *copula* (as opposed to the survival copula) of a d -dimensional distribution \mathbf{Y} under consideration, then $C^n(\mathbf{u}^{1/n})$ will be the copulas of the n -block *maxima*, defined as

$$\mathbf{M}_n = (M_{1,n}, \dots, M_{d,n}) := (\max\{Y_{1,k}\}_{k=1}^n, \dots, \max\{Y_{d,k}\}_{k=1}^n),$$

where $\mathbf{Y}_n = (Y_{1,n}, \dots, Y_{d,n})$, $n \geq 1$, form a sequence of independent copies of \mathbf{Y} . Indeed, if G is the joint distribution function of \mathbf{Y} , independence of the \mathbf{M}_n implies (cf. the proof of Lemma 6.1) that

$$\mathbb{P}[\mathbf{M}_n \leq \mathbf{x}] = (\mathbb{P}[\mathbf{Y} \leq \mathbf{x}])^n = G^n(\mathbf{x}).$$

Hence the marginals are $\mathbb{P}[M_{i,n} \leq x_i] = G_i^n(x_i)$, and an explicit computation yields the form $C^n(\mathbf{u}^{1/n})$ of the copula of \mathbf{M}_n . Corollary 6.4 of course applies also in this case. To summarize, we have the following (as before, C is Archimedean with generator ψ):

- (i) If C is the *copula* of the distribution \mathbf{Y} , then the dependence structure of the corresponding n -block *maxima* can be approximated by the Gumbel copula for large n (if $1 - \psi(x^{-1}) \in \text{RV}_{-\alpha}$).
- (ii) If C is the *survival copula* of the distribution \mathbf{Y} , then the dependence structure of the corresponding n -block *minima* can be approximated by the Gumbel copula for large n (if $1 - \psi(x^{-1}) \in \text{RV}_{-\alpha}$).

We finally remark that Proposition 4.11 allows us to express the condition on ψ in terms of a condition on the corresponding radial part R —namely, $1/R$ must have a regularly varying tail.

6.2 Componentwise maxima of ℓ_1 -norm symmetric distributions

In this section, \mathbf{M}_n will denote the componentwise n -block maxima of the i.i.d. ℓ_1 -norm symmetric random vectors $\mathbf{X}_1, \mathbf{X}_2, \dots$, i.e.

$$\mathbf{M}_n = (M_{1,n}, \dots, M_{d,n}) := (\max\{X_{1,k}\}_{k=1}^n, \dots, \max\{X_{d,k}\}_{k=1}^n).$$

As before, $R \sim F$ with $R > 0$ a.s. is the radial part of \mathbf{X}_n , and $\psi = \mathfrak{W}_d F$ is the corresponding Archimedean generator. The generated survival copula is denoted C . Unlike the minima \mathbf{W}_n considered in Section 6.1, the \mathbf{M}_n are *not* ℓ_1 -norm symmetric. We thus need to take a different approach to analyze their asymptotic behavior, and it turns out that the theory of multivariate regular variation and multivariate extremes provides the proper framework. The stochastic representation of the ℓ_1 -norm symmetric distributions renders the computations particularly simple.

Since the multivariate extreme value theory is more complicated than its univariate counterpart, and because it is only used in this section, we refrain from giving an introduction to the general theory. Instead we give precise references for results that are used. The book [25] by Resnick contains all theory we need here, and much more beyond that.

Proposition 6.5 *Suppose \mathbf{X} is ℓ_1 -norm symmetric with radial part $R \in \text{RV}_{-\alpha}$, $\alpha \in (0, \infty)$. Then \mathbf{X} is in the domain of attraction of*

$$G(\mathbf{x}) := \exp\{-\mu([\mathbf{0}, \mathbf{x}]^c)\}, \quad \mathbf{x} > \mathbf{0}.$$

where the exponent measure μ is given by

$$\mu\{\mathbf{x} \in \mathbb{R}_+^d \setminus \{\mathbf{0}\} \mid \|\mathbf{x}\|_1 > r, \mathbf{x}/\|\mathbf{x}\|_1 \in A\} = r^{-\alpha} \sigma(A), \quad r > 0, A \in \mathcal{S}^d,$$

and the spectral measure σ is the uniform distribution on \mathcal{S}^d .

Proof. Since \mathbf{X} is ℓ_1 -norm symmetric, Proposition 2.9 implies that $\mathbf{S} \stackrel{\text{d}}{=} \mathbf{X}/\|\mathbf{X}\|_1$ is uniformly distributed on \mathcal{S}^d , $\|\mathbf{X}\|_1 \stackrel{\text{d}}{=} R$, and $\|\mathbf{X}\|_1$ and $\mathbf{X}/\|\mathbf{X}\|_1$ are independent. Hence for any Borel subset $A \subseteq \mathcal{S}^d$ and $x > 0$,

$$\lim_{u \rightarrow \infty} \frac{\mathbb{P}[\|\mathbf{X}\|_1 > ux, \mathbf{X}/\|\mathbf{X}\|_1 \in A]}{\mathbb{P}[\|\mathbf{X}\|_1 > u]} = \lim_{u \rightarrow \infty} \frac{\bar{F}(ux)}{\bar{F}(u)} \mathbb{P}[\mathbf{S} \in A] = x^{-\alpha} \mathbb{P}[\mathbf{S} \in A].$$

Setting $t := 1/\bar{F}(u)$ and $a_t := (1/\bar{F})^\leftarrow(t)$, we also have

$$\lim_{u \rightarrow \infty} \frac{\mathbb{P}[\|\mathbf{X}\|_1 > ux, \mathbf{X}/\|\mathbf{X}\|_1 \in A]}{\mathbb{P}[\|\mathbf{X}\|_1 > u]} = \lim_{t \rightarrow \infty} t \mathbb{P}[a_t^{-1} \|\mathbf{X}\|_1 > \mathbf{x}, \mathbf{X}/\|\mathbf{X}\|_1 \in A].$$

We can now apply Corollary 5.18 (b) in Resnick [25] to reach the desired conclusion with $\sigma = \mathbb{P}[\mathbf{S} \in (\cdot)]$. \square

This means in particular that \mathbf{X} is *multivariate regularly varying* with tail index α and spectral measure σ . We now express the exponent measure μ directly in terms of σ .

Lemma 6.6 *Let μ and σ be as in Proposition 6.5. Then*

$$\mu([\mathbf{0}, \mathbf{x}]^c) = \int_{\mathcal{S}^d} \max \left\{ \frac{w_1^\alpha}{x_1^\alpha}, \dots, \frac{w_d^\alpha}{x_d^\alpha} \right\} \sigma(d\mathbf{w}), \quad \mathbf{x} \in \mathbb{R}_+^d.$$

Proof. We slightly modify the argument in Resnick [25], p. 267-268, to account for the fact that α need not be one.

Define the polar coordinate transformation $T : \mathbf{x} \mapsto (\|\mathbf{x}\|_1, \mathbf{x}/\|\mathbf{x}\|_1)$. Then by Proposition 6.5, $\mu \circ T^{-1}(r, A) = r^{-\alpha} \sigma(A)$ for $r > 0$, $A \subseteq \mathcal{S}^d$, with density $\alpha r^{-\alpha-1} dr \sigma(d\mathbf{w})$. Of course $\mu([\mathbf{0}, \mathbf{x}]^c) = \mu \circ T^{-1} \circ T([\mathbf{0}, \mathbf{x}]^c)$, and the same computations as in [25], p. 267 yield

$$T([\mathbf{0}, \mathbf{x}]^c) = \left\{ (r, \mathbf{w}) \mid r > \min \left\{ \frac{x_1}{w_1}, \dots, \frac{x_d}{w_d} \right\} \right\}.$$

We can now integrate the density of $\mu \circ T^{-1}$ over this region:

$$\begin{aligned} \mu([\mathbf{0}, \mathbf{x}]^c) &= \int \int_{T([\mathbf{0}, \mathbf{x}]^c)} \alpha r^{-\alpha-1} dr \sigma(d\mathbf{w}) \\ &= \int_{\mathcal{S}^d} \int_{\left\{ r > \min \left\{ \frac{x_1}{w_1}, \dots, \frac{x_d}{w_d} \right\} \right\}} \alpha r^{-\alpha-1} dr \sigma(d\mathbf{w}) \\ &= \int_{\mathcal{S}^d} \left[\min \left\{ \frac{x_1}{w_1}, \dots, \frac{x_d}{w_d} \right\} \right]^{-\alpha} \sigma(d\mathbf{w}), \end{aligned}$$

which is equal to the asserted expression. \square

It is now possible to compute the copula C_0 of the limiting distribution G . Although it does not have a particularly elegant form (except in the bivariate case, as we shall see), it does have an important interpretation: if C is the copula of the distribution \mathbf{Y} , then corresponding n -block *minima* have a dependence structure approximated by C_0 for large n . Equivalently, C_0 describes the dependence of the n -block *maxima* of \mathbf{Y} in case C is the *survival copula* of \mathbf{Y} . See the discussion at the end of Section 6.1.

First, to obtain the marginals G_1 of G (they are all equal due to exchangeability) we note that by Proposition 6.5, $G_1(x) = \exp\{-\mu(B_x)\}$, where $B_x := [\mathbf{0}, (x, \infty, \dots, \infty)]^c$. Lemma 6.6 then gives

$$\mu(B_x) = x^{-\alpha} \int_{\mathcal{S}^d} w_1^\alpha \sigma(d\mathbf{w}) = c_\alpha x^{-\alpha},$$

with $c_\alpha = \int_{\mathcal{S}^d} w_1^\alpha \sigma(d\mathbf{w})$. Hence $G_1(x) = \exp\{-c_\alpha x^{-\alpha}\}$, which in particular means that G_1 is of the same type as Φ_α . This is to be expected since the marginal survival function of \mathbf{X} is $\psi = \mathfrak{W}_d F$, which is in $\text{MDA}(\Phi_\alpha)$ by Proposition 4.3 and Theorem 3.5 (we still work under the assumption of Proposition 6.5: $R \in \text{RV}_{-\alpha}$ with $\alpha \in (0, \infty)$).

Inverting G_1 gives $G_1^\leftarrow(u) = (-c_\alpha^{-1} \log u)^{-1/\alpha}$, and we may apply Lemma 6.6 to compute the copula of G as $C_0(\mathbf{u}) = G(G_1^{-1}(u_1), \dots, G_1^{-1}(u_d))$:

$$\begin{aligned} C_0(\mathbf{u}) &= \exp \left\{ - \int_{\mathcal{S}^d} \max \left\{ -\frac{\log u_1}{c_\alpha} w_1^\alpha, \dots, -\frac{\log u_d}{c_\alpha} w_d^\alpha \right\} \sigma(d\mathbf{w}) \right\} \\ &= \exp \left\{ -\frac{1}{c_\alpha} \int_{\mathcal{S}^d} \max \{ -w_1^\alpha \log u_1, \dots, -w_d^\alpha \log u_d \} \sigma(d\mathbf{w}) \right\}. \end{aligned}$$

Although by Sklar's Theorem we already know that C_0 is a copula, one may readily verify that the marginals are standard uniform. It is also easy to check that C_0 satisfies the EV copula property $C_0^n(\mathbf{u}^{1/n}) = C_0(\mathbf{u})$.

Example 6.7 For $d = 2$, C_0 can be computed explicitly due to the simple form of the spectral measure σ . The first step is to calculate $V(x_1, x_2) := \int_{\mathcal{S}^2} \max\{w_1^\alpha x_1, w_2^\alpha x_2\} \sigma(d\mathbf{w})$ for $x_1, x_2 > 0$. We parameterize \mathcal{S}^2 by $t := w_1$ and obtain

$$V(x_1, x_2) = \int_0^1 \max\{x_1 t^\alpha, x_2(1-t)^\alpha\} dt = \int_0^\kappa x_2(1-t)^\alpha dt + \int_\kappa^1 x_1 t^\alpha dt,$$

where $\kappa = x_2^{1/\alpha} / (x_1^{1/\alpha} + x_2^{1/\alpha})$ solves the equation $x_1 \kappa^\alpha = x_2(1-\kappa)^\alpha$. The two integrals on the right side are straightforward to evaluate, and the result is

$$V(x_1, x_2) = \frac{1}{1+\alpha} \left(x_1 + x_2 - \left(x_1^{-1/\alpha} + x_2^{-1/\alpha} \right)^{-\alpha} \right).$$

The quantity c_α is obtained similarly: $c_\alpha = \int_{\mathcal{S}^d} w_1^\alpha \sigma(d\mathbf{w}) = \int_0^1 t^\alpha dt = (1+\alpha)^{-1}$. Piecing things together we get $C_0(u_1, u_2) = \exp\{-(1+\alpha)V(-\log u_1, -\log u_2)\}$, or, more explicitly,

$$C_0(u_1, u_2) = u_1 u_2 \exp \left\{ \left((-\log u_1)^{-1/\alpha} + (-\log u_2)^{-1/\alpha} \right)^{-\alpha} \right\}.$$

This is nothing but the well-known Galambos copula.

We conclude this section with a remark regarding the fact that we only considered the heavy-tailed case, i.e. $R \in \text{RV}_{-\alpha}$ for $\alpha \in (0, \infty)$, or, equivalently, R being in the Fréchet domain of attraction. What would happen if R were in the Gumbel domain? Assuming unbounded support for simplicity we have

$$\lim_{u \rightarrow \infty} \frac{\mathbb{P}[\|\mathbf{X}\|_1 > ux, \mathbf{X}/\|\mathbf{X}\|_1 \in A]}{\mathbb{P}[\|\mathbf{X}\|_1 > u]} = \lim_{u \rightarrow \infty} \frac{\bar{F}(ux)}{\bar{F}(u)} \mathbb{P}[\mathbf{S} \in A] = 0$$

for all Borel sets $A \subseteq \mathcal{S}^d$. The limiting measure μ in Proposition 6.5 would thus be identically zero, $\mu([\mathbf{0}, \mathbf{x}]^c) = 0$ for $\mathbf{x} \in \mathbb{R}_+^d \setminus \{\mathbf{0}\}$, meaning that the limit distribution would be degenerate with a point mass at the origin. Although a different normalization (such as $(\cdot - u)/a(u)$ as in Theorem 3.5) might still yield meaningful results, we do not pursue this here since it involves translations and thus is non-standard in the basic theory of multivariate regular variation and multivariate extremes.

Chapter 7

Archimedean threshold copulas

In the univariate theory of extremes, the problem of determining the limiting distributions of maxima (or minima) is intimately related to that of finding the limiting distributions of threshold exceedances. In the multivariate setting things become more complicated, since thresholds can have many different forms and diverge in many different ways. In this chapter we consider the limiting dependence structures of *componentwise* exceedances when the underlying distribution has an Archimedean copula. The crucial tool will, as before, be the ℓ_1 -norm symmetric distributions.

7.1 Lower threshold copulas

Let C be a d -dimensional copula and suppose (U_1, \dots, U_d) is distributed according to C . For given quantile levels $\alpha_i \in (0, 1]$,

$$\mathbb{P}[U_1 \leq \beta_1, \dots, U_d \leq \beta_d \mid U_1 \leq \alpha_1, \dots, U_d \leq \alpha_d], \quad \beta_i \in [0, \alpha_i], \quad (7.1)$$

defines a distribution function on $[0, \alpha_1] \times \dots \times [0, \alpha_d]$. The copula of this distribution is called the *lower threshold copula* of C ; the limit as each α_i decreases is called the *limiting lower threshold copula* (LLTC). If C is Archimedean, the LLTC is always the Clayton copula, whose generator is $\psi_\theta^{\text{Cl}}(x) = (1 + \theta x)^{-1/\theta}$. This result follows in a very transparent way using the connection with ℓ_1 -norm symmetric distributions.

We need to be more precise about the statement “each α_i decreases”. It can happen that the set $L_0 := \{\mathbf{u} \in [0, 1]^d \mid C(\mathbf{u}) = 0\}$, called the *zero set* of C , has positive measure (w.r.t. Lebesgue measure on $[0, 1]^d$.) If C is Archimedean, this happens when the generator ψ is non-strict. In this case one cannot let $\alpha_i \downarrow 0$, since eventually the probabilities in (7.1) will not be well-defined. Instead, we should require $\boldsymbol{\alpha} := (\alpha_1, \dots, \alpha_d) \notin L_0$ and $\boldsymbol{\alpha} \rightarrow \mathbf{u}_0$ for some $\mathbf{u}_0 \in \partial L_0$. If $L_0 \neq \{\mathbf{0}\}$ there are different possible choices for \mathbf{u}_0 , and even when $L_0 = \{\mathbf{0}\}$ there are no guarantees a priori that the LLTC is independent of how $\boldsymbol{\alpha}$ approaches $\mathbf{u}_0 = \mathbf{0}$. To avoid ambiguities, one therefore usually

takes $\alpha_1 = \alpha_2 = \dots = \alpha_d$ and lets α_i decrease until it hits ∂L_0 (this is the case in e.g. McNeil et al. [20] and Juri and Wüthrich [18]). Since in the Archimedean case, $L_0 = \{\mathbf{u} \in [0, 1]^d \mid \sum_{i=1}^d \psi^{-1}(u_i) \geq \psi^{-1}(0)\}$, we are able to handle this complication by supposing that $\sum_{i=1}^d \psi^{-1}(\alpha_i) \uparrow \psi^{-1}(0)$, which ensures that $\boldsymbol{\alpha}$ approaches L_0 from above. Moreover, it turns out that neither the relative rates of convergence of the α_i , nor which $\mathbf{u}_0 \in L_0$ they tend to, affects the LLTC. We therefore allow the quantiles α_i to be different in order to emphasize this fact.

Now, suppose C is Archimedean with generator ψ . Since ψ^{-1} is decreasing, $\{U_1 \leq u\} = \{\psi^{-1}(U_1) \geq \psi^{-1}(u)\}$ for $u \in [0, 1]$. With $X_i := \psi^{-1}(U_i)$, $v_i := \psi^{-1}(\alpha_i)$ and $x_i := \psi^{-1}(\beta_i) - \psi^{-1}(\alpha_i) \geq 0$, (7.1) can thus be written as

$$\begin{aligned} \mathbb{P}[U_1 \leq \beta_1, \dots, U_d \leq \beta_d \mid U_1 \leq \alpha_1, \dots, U_d \leq \alpha_d] &= \\ &= \mathbb{P}[X_1 - v_1 \geq x_1, \dots, X_d - v_d \geq x_d \mid X_1 \geq v_1, \dots, X_d \geq v_d] \\ &=: \bar{H}_{\mathbf{v}}(\mathbf{x}), \end{aligned} \quad (7.2)$$

where $\mathbf{x} = (x_1, \dots, x_d)$ with $x_i \in [0, \psi^{-1}(0) - v_i]$ and $\mathbf{v} = (v_1, \dots, v_d)$ denotes the vector of thresholds. Since all X_i are continuous, the inequalities can be taken as strict, and we see that $\bar{H}_{\mathbf{v}}$ defines a survival function on the interior of \mathbb{R}_+^d .

Lemma 7.1 *$\mathbf{X} = (X_1, \dots, X_d)$ as defined above is ℓ_1 -norm symmetric with survival copula generator ψ . Moreover, $\mathbf{X}_{\mathbf{v}} \sim \bar{H}_{\mathbf{v}}$ is also ℓ_1 -norm symmetric with survival copula generator $\psi_{\mathbf{v}}$ given by*

$$\psi_{\mathbf{v}}(x) = \frac{\psi(x + \|\mathbf{v}\|_1)}{\psi(\|\mathbf{v}\|_1)}.$$

Proof. The statement about \mathbf{X} follows from Corollary 2.13. An explicit computation yields

$$\bar{H}_{\mathbf{v}}(\mathbf{x}) = \frac{\mathbb{P}[X_1 - v_1 > x_1, \dots, X_d - v_d > x_d]}{\mathbb{P}[X_1 > v_1, \dots, X_d > v_d]} = \frac{\psi(\|\mathbf{x}\|_1 + \|\mathbf{v}\|_1)}{\psi(\|\mathbf{v}\|_1)},$$

whence the last assertion due to Theorem 2.12. \square

Due to (7.2), the survival copula corresponding to $\bar{H}_{\mathbf{v}}$, which we denote by $C_{\mathbf{v}}$, is the lower threshold copula of C . Now, $\alpha_i \downarrow 0$ if and only if $v_i \uparrow \psi^{-1}(0)$, so in view of the discussion in the beginning of this section the LLTC of C is given by the limit of $C_{\mathbf{v}}$ as $\|\mathbf{v}\|_1 \uparrow \psi^{-1}(0)$.

Proposition 7.2 *Let ψ and $\mathbf{X}_{\mathbf{v}}$ be as in Lemma 7.1, and suppose ψ is in some domain of attraction. Then there exists a normalization $\beta(\cdot)$ such that $\mathbf{X}_{\mathbf{v}}/\beta(\|\mathbf{v}\|_1)$ converges weakly to an ℓ_1 -norm symmetric distribution with survival copula generator $\psi_{\theta}^{\text{Cl}}$ as $\|\mathbf{v}\|_1 \rightarrow \psi^{-1}(0)$. The parameter θ is given by:*

- (i) *If ψ is in the Fréchet domain, i.e. $\psi(x) \in \text{RV}_{-\alpha}$, for some $\alpha \in (0, \infty)$, then $\theta = 1/\alpha$.*

(ii) If ψ is in the Gumbel domain, then $\theta = 0$.

(iii) If ψ is in the Weibull domain, i.e. $\psi(x^* - x^{-1}) \in \text{RV}_{-\alpha}$ for some $\alpha \in (0, \infty)$ and where $x^* = \psi^{-1}(0)$, then $\theta = -1/\alpha$.

Proof. In the Fréchet case Lemma 7.1 and the fact that $\psi \in \text{RV}_{-\alpha}$ yield

$$\mathbb{P} \left[\frac{\mathbf{X}_v}{\|\mathbf{v}\|_1} > \mathbf{x} \right] = \frac{\psi(\|\mathbf{v}\|_1(1 + \|\mathbf{x}\|_1))}{\psi(\|\mathbf{v}\|_1)} \rightarrow (1 + \|\mathbf{x}\|_1)^{-\alpha} = \psi_{1/\alpha}^{\text{Cl}}(\|\mathbf{x}\|_1)$$

as $\|\mathbf{v}\|_1 \rightarrow \psi^{-1}(0) = \infty$. In this case we may thus choose $\beta(x) = x$. In the Weibull case, let t be defined via $\|\mathbf{v}\|_1 = x^* - t^{-1}$, and set $\beta(x) := x^* - x$. Then

$$\begin{aligned} \mathbb{P} \left[\frac{\mathbf{X}_v}{\beta(\|\mathbf{v}\|_1)} > \mathbf{x} \right] &= \frac{\psi(\|\mathbf{v}\|_1 + \beta(\|\mathbf{v}\|_1)\|\mathbf{x}\|_1)}{\psi(\|\mathbf{v}\|_1)} \\ &= \frac{\psi(x^* - (1 - \|\mathbf{x}\|_1)t^{-1})}{\psi(x^* - t^{-1})} \\ &\rightarrow (1 - \|\mathbf{x}\|_1)^\alpha = \psi_{-1/\alpha}^{\text{Cl}}(\|\mathbf{x}\|_1), \quad \|\mathbf{v}\|_1 \rightarrow x^*, \end{aligned}$$

since $\psi(x^* - x^{-1}) \in \text{RV}_{-\alpha}$. In the Gumbel case we take β to be the auxiliary function of ψ . The right side of the first line in the previous display now tends to $\exp\{-\|\mathbf{x}\|_1\} = \psi_0^{\text{Cl}}(\|\mathbf{x}\|_1)$ due to the property (3.2) in Theorem 3.5. In each of the three cases, the assertion that the limit is ℓ_1 -norm symmetric follows from Theorem 2.12. \square

Remark. The function $\beta(\cdot)$ has the standard interpretation as the proper normalization for the conditional univariate exceedances $(X_1 - v | X_1 > v)$ to converge to a Generalized Pareto Distribution (cf. Balkema and de Haan [1] and Pickands [23]).

Note that the parameter θ can be negative. This happens when $a := \psi^{-1}(0) < \infty$ and $1 - \psi(a - x^{-1}) \in \text{RV}_{-\alpha}$, $\alpha > 0$. From Proposition 4.9 we know that one has the limitation $\theta = -1/\alpha \geq -1/(d-1)$, which is consistent with the range of parameter values for which the Clayton copula is defined (see McNeil and Nešlehová [19], Example 4.)

Corollary 7.3 *Let C be an Archimedean copula whose generator ψ is in some domain of attraction. The corresponding limiting lower threshold copula (LLTC) is then Clayton with parameter θ as in Proposition 7.2.*

Proof. Denoting again the copula if \bar{H}_v by C_v , the LLTC is given by the pointwise limit of C_v as $\|\mathbf{v}\|_1 \rightarrow \psi^{-1}(0)$. But scaling by a positive constant does not change the copula¹, so $\mathbf{X}_v/\beta(\|\mathbf{v}\|_1)$, like \mathbf{X}_v , has survival copula C_v . Since weak convergence of the distributions implies pointwise convergence of the copulas (cf. Feidt et al. [12]), the assertion follows from Proposition 7.2. \square

¹In fact, copulas are invariant under strictly increasing transformations T_i of the components of the underlying random vector—if $(Y_1, Y_2) \sim H$ with marginals H_1, H_2 has copula C , then the copula of $(T_1(Y_1), T_2(Y_2))$ is $H(T_1^{\leftarrow} \circ T_1 \circ H_1^{\leftarrow}(\cdot), T_2^{\leftarrow} \circ T_2 \circ H_2^{\leftarrow}(\cdot)) = C$ by the properties of generalized inverses.

The fact that only the sum $\|\mathbf{v}\|_1$ has to tend to $\psi^{-1}(0)$ immediately gives the form of all *marginal LLTCs*, which are defined as the limits of

$$\mathbb{P}[U_i \leq \beta_i, i \in J \mid U_i \leq \alpha_i, i \in J], \quad \beta_i \in [0, \alpha_i], \quad (7.3)$$

as $(\alpha_i)_{i \in J}$ tends to the corresponding zero set, where $J \subset \{1, \dots, d\}$ is a non-empty index set. More precisely, we have

Corollary 7.4 *Let C be an Archimedean copula whose generator ψ is in some domain of attraction. All corresponding marginal LLTCs are then Clayton with parameter θ as in Proposition 7.2.*

Proof. (7.3) is obtained from (7.1) upon setting $\alpha_i = 1$ and $\beta_i = 1$ for all $i \notin J$. This corresponds to $x_i = v_i = 0$, $i \notin J$. Letting the remaining v_i increase such that $\|\mathbf{v}\|_1 \uparrow \psi^{-1}(0)$, the assertion is inferred from Proposition 7.2 as in Corollary 7.3. \square

In a modeling situation, there might be given requirements on the behavior of the LLTC of the copula C . If one is working with the stochastic representation in terms of ℓ_1 -norm symmetric distributions, the results in Chapter 4 translate the three conditions on ψ given in Proposition 7.2 into conditions on the corresponding radial part R . It may also be the case that a certain R has been chosen through some inference procedure, and one then wants to know the implications for the extremal behavior of the limiting copula. Again, Chapter 4 provides the necessary link.

We conclude with a remark on the connection between threshold exceedances and the componentwise maxima treated in Section 6.2. One may view the conclusion of Example 6.7 as an extension of the well-known fact that the Galambos copula is the EV limit of the survival copula of the bivariate Pareto distribution (see McNeil et al. [20], Example 7.50). In fact, the (symmetric) bivariate Pareto is ℓ_1 -norm symmetric and its survival copula is Clayton (this is easily checked: its joint survival function is $\bar{H}(x_1, x_2) = (x_1 + x_2 + 1)^{-\alpha}$ with $\alpha > 0$).

Now, by Proposition 7.2, any heavy-tailed bivariate ℓ_1 -norm symmetric distribution has asymptotically the same shape as the bivariate Pareto far out in the tail. In the case of ℓ_1 -norm symmetric distributions, therefore, the following recipe produces correct results: given $\mathbf{X} = (X_1, X_2)$, compute the limiting distribution \mathbf{X}_∞ of the threshold exceedances. Take the componentwise maxima of \mathbf{X}_∞ and compute the corresponding limiting distribution. The copula of this limit is the same as the copula of the distribution that is obtained by directly taking the limit of the componentwise maxima of \mathbf{X} . Intuitively such a result is reasonable; for large n , the n -block maxima will be located far out in the tail, where the underlying distribution looks approximately like the limit of the exceedances.

The last statement gives rise to an interesting question: in general, what are the conditions on the underlying multivariate distribution for these heuristics to be correct? I.e., under which circumstances does the recipe described above give correct results?

7.2 Upper threshold copulas

In analogy with the lower threshold copulas, one may define the *upper threshold copula* for given quantiles $\alpha_1, \dots, \alpha_d$ as the copula of the distribution

$$\mathbb{P}[U_1 > \beta_1, \dots, U_d > \beta_d \mid U_1 > \alpha_1, \dots, U_d > \alpha_d], \quad \beta_i \in [\alpha_i, 1], \quad (7.4)$$

where again (U_1, \dots, U_d) are distributed according to the copula C . The limit as the $\alpha_i \uparrow 1$ is called the *limiting upper threshold copula* (LUTC)², and our aim is to investigate the extent to which our previous methodology can be applied to compute the LUTC in the case when C is Archimedean.

We first define $X_i := \psi^{-1}(U_i)$, $v_i := \psi^{-1}(\alpha_i)$ and $x_i := \psi^{-1}(\beta_i)$. The quantity (7.4) can then be expressed in terms of the ℓ_1 -norm symmetric random variable $\mathbf{X} = (X_1, \dots, X_d)$, whose survival copula is C (that \mathbf{X} is ℓ_1 -norm symmetric follows as before from Corollary 2.13):

$$\begin{aligned} \mathbb{P}[U_1 > \beta_1, \dots, U_d > \beta_d \mid U_1 > \alpha_1, \dots, U_d > \alpha_d] &= \\ &= \mathbb{P}[\mathbf{X} \leq \mathbf{x} \mid \mathbf{X} \leq \mathbf{v}] \\ &=: H_{\mathbf{v}}(\mathbf{x}), \quad x_i \in [0, v_i]. \end{aligned}$$

To get the inequalities right we used strict monotonicity of ψ^{-1} and the fact that \mathbf{X} has continuous marginals. It is straightforward to show that $H_{\mathbf{v}}$ is *not* the distribution function of a ℓ_1 -norm symmetric distribution—indeed, the support of $(\mathbf{X} \mid \mathbf{X} \leq \mathbf{v})$ is $[\mathbf{0}, \mathbf{v}]$, which is not consistent with ℓ_1 -norm symmetry. Its survival copula is therefore not Archimedean, despite the fact that C is. Hence the methodology used in the LLTC case cannot be employed, and this difficulty cannot be overcome using basic multivariate extreme value theory as easily as in the EV copula case in Section 6.2. In fact, the ULTC cannot be explicitly computed, but we shall derive a distribution function whose survival copula is the ULTC of C . For simplicity we carry out the calculations in two dimensions to illustrate the procedure. Generalizing the results to higher dimension is straightforward.

With $d = 2$, we have

$$\begin{aligned} \mathbb{P}[X_1 \leq x_1, X_2 \leq x_2] &= 1 - \mathbb{P}[X_1 > x_1] - \mathbb{P}[X_2 > x_2] + \mathbb{P}[X_1 > x_1, X_2 > x_2] = \\ &= 1 - \psi(x_1) - \psi(x_2) + \psi(x_1 + x_2) = \\ &= G(x_1) + G(x_2) - G(x_1 + x_2), \end{aligned}$$

where $G(x) := 1 - \psi(x)$ is the marginal distribution function of \mathbf{X} . In the general case one may use the sieve formula to express $\mathbb{P}[\mathbf{X} \leq \mathbf{x}]$ in terms of $G(\sum_{i \in J} x_i)$ for subsets $J \subseteq \{1, \dots, d\}$. Proceeding with the case $d = 2$ we obtain

$$H_{\mathbf{v}}(\mathbf{x}) = \frac{G(x_1) + G(x_2) - G(x_1 + x_2)}{G(v_1) + G(v_2) - G(v_1 + v_2)}, \quad x_i \in [0, v_i].$$

²We do not have the same problem as in the LLTC case—the set $L_1 := \{\mathbf{u} \in [0, 1]^d \mid C(\mathbf{u}) = 1\}$ is always a nullset. Indeed, $C(u_1, \dots, u_d) \leq C(u_1, 1, \dots, 1) = u_1$ implies that $C(\mathbf{u}) < 1$ for $\mathbf{u} \neq \mathbf{1}$.

Now, $\alpha_i \uparrow 1$ is equivalent to $v_i \downarrow 0$, so in order to have a non-degenerate limit we define $\tilde{\mathbf{X}} = (\tilde{X}_1, \tilde{X}_2)$ as the normalized $[0, 1]^2$ -valued random variable (depending on \mathbf{v}) distributed according to

$$\mathbb{P}[\tilde{X}_1 \leq x_1, \tilde{X}_2 \leq x_2] = H_{\mathbf{v}}(x_1 v_1, x_2 v_2), \quad x_i \in [0, 1].$$

The copula is not affected, since it is invariant under strictly increasing transformations of the components. We may therefore consider the copula of $\tilde{\mathbf{X}}$ as v_1 and v_2 tend to zero.

Proposition 7.5 *Suppose $1/X_1 \sim \psi(x^{-1})$ is in some domain of attraction, and let $v_1 = v_2 = v$. In particular we then have $1 - \psi(x^{-1}) \in RV_{-\alpha}$ for some $\alpha \in (0, 1]$, and if in addition $\alpha < 1$, then*

$$H_0(x_1, x_2) := \lim_{v \downarrow 0} \mathbb{P}[\tilde{X}_1 \leq x_1, \tilde{X}_2 \leq x_2] = \frac{x_1^\alpha + x_2^\alpha - (x_1 + x_2)^\alpha}{2 - 2^\alpha}.$$

Proof. Theorem 3.5 implies that $G(x^{-1}) = 1 - \psi(x^{-1}) \in RV_\alpha(0)$ for some $\alpha \in (0, \infty]$, and by Proposition 4.12, $\alpha \leq 1$. Since $v_1 = v_2 = v$ this yields

$$\begin{aligned} \mathbb{P}[\tilde{X}_1 \leq x_1, \tilde{X}_2 \leq x_2] &= \frac{G(x_1 v) + G(x_2 v) - G(x_1 v + x_2 v)}{2G(v) - G(2v)} = \\ &= \frac{\frac{G(x_1 v)}{G(v)} + \frac{G(x_2 v)}{G(v)} - \frac{G(v(x_1 + x_2))}{G(v)}}{2 - \frac{G(2v)}{G(v)}} \\ &\rightarrow \frac{x_1^\alpha + x_2^\alpha - (x_1 + x_2)^\alpha}{2 - 2^\alpha}, \quad v \downarrow 0. \end{aligned}$$

□

The assumption $v_1 = v_2$ is not restrictive if we are only interested in the copulas. Indeed, if $v_2/v_1 \rightarrow c > 0$, then one can show, using that the convergence in Proposition 7.5 is locally uniform (a basic result from Karamata theory), that $H_0(x_1, cx_2)$ is obtained in the limit. This just represents a linear scaling of the second component, which leaves the copula unchanged. The marginals $H_0(x, 1) = (x^\alpha + 1 - (x + 1)^\alpha)/(2 - 2^\alpha)$ do not have closed form inverses, so we are not able to explicitly compute the copula or the survival copula. As a final remark we mention that more precise results can be obtained in the limiting case when $\alpha = 1$, see Charpentier and Segers [6].

Bibliography

- [1] A. A. Balkema and L. de Haan. Residual life time at great age. *Ann. Probability*, 2:792–804, 1974.
- [2] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.
- [3] L. Breiman. On some limit theorems similar to the arc-sin law. *Theory of Probability and its Applications*, X:323–331, 1965.
- [4] L. Breiman. *Probability*. Addison-Wesley, Reading, MA, 1968.
- [5] Philippe Capéraà, Anne-Laure Fougères, and Christian Genest. Bivariate distributions with given extreme value attractor. *J. Multivariate Anal.*, 72(1):30–49, 2000.
- [6] A. Charpentier and J. Segers. Tails of multivariate archimedean copulas. Unpublished preprint, 2008.
- [7] S. Coles, J. Heffernan, and Tawn J. Dependence measures for extreme value analyses. *Extremes*, 2:4:339–365, 2000.
- [8] S. Demarta. *The Copula Approach to Modelling Multivariate Extreme Values*. Dissertation ETH No. 17307, 2007.
- [9] P. Embrechts, A. J. McNeil, and D. Straumann. Correlation and dependency in risk management: Properties and pitfalls. *Risk Management: Value at Risk and Beyond*, pages 176–223, 2002.
- [10] Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modelling extremal events*, volume 33 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997. For insurance and finance.
- [11] K.-T. Fang, S. Kotz, and K.-W. Ng. *Symmetric Multivariate and Related Distributions*. Chapman & Hall, New York, 1990.
- [12] A. Feidt, C. Genest, and J. Nešlehová. Asymptotics of joint maxima for discontinuous random variables. Submitted, 2008.

- [13] R.A. Fisher and L.H.C. Tippett. Limiting forms of the frequency distributions of largest or smallest member of a sample. *Proc. of Cambridge Philos. Soc.*, 24:180–190, 1928.
- [14] J. Galambos. Order statistics of samples from multivariate distributions. *Journal of the American Statistical Association*, 70:674–680, 1975.
- [15] C. Genest and L. Rivest. A characterization of gumbel’s family of extreme value distributions. *Statistics & Probability Letters*, 8:207–211, 1989.
- [16] B. Gnedenko. Sur la distribution limite du terme maximum d’une série aléatoire. *Ann. of Math. (2)*, 44:423–453, 1943.
- [17] L. de Haan. A form of regular variation and its application to the domain of attraction of the double exponential distribution. *Wahrscheinlichkeitstheorie verw. Geb.*, 17:241–258, 1971.
- [18] Alessandro Juri and Mario V. Wüthrich. Copula convergence theorems for tail events. *Insurance Math. Econom.*, 30(3):405–420, 2002.
- [19] A.J. McNeil and J. Nešlehová. Multivariate archimedean copulas, d -monotone functions and l_1 -norm symmetric distributions. *Ann. Stat.*, to appear, 2008.
- [20] Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, Princeton, NJ, 2005.
- [21] Thomas Mikosch. Copulas: Tales and facts (with discussion). *Extremes*, 9:1–62, 2007.
- [22] Roger B. Nelsen. *An Introduction to Copulas*, volume 139. Springer, New York, second edition, 2006.
- [23] James Pickands, III. Statistical inference using extreme order statistics. *Ann. Statist.*, 3:119–131, 1975.
- [24] M. Proschan and B. Presnell. Expect the unexpected from conditional expectation. *The American Statistician*, 52:248–252, 1998.
- [25] Sidney I. Resnick. *Extreme values, regular variation and point processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. Reprint of the 1987 original.
- [26] Abe Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231, 1959.
- [27] R. E. Williamson. Multiply monotone functions and their Laplace transforms. *Duke Math. J.*, 23:189–207, 1956.